High-Throughput Study of Anti-Glycan Antibodies in Serum of Women with Epithelial Ovarian Cancer: Biomarker Detection and Investigation of Glycan-Binding Patterns

Dissertation
zur
Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)
vorgelegt der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Universität Zürich
von
Francis Jacob
aus
Deutschland

<u>Promotionskomitee</u>

Prof. Dr.	Christian von Mering	(Vorsitz)
PD Dr.	Viola Heinzelmann-Schwarz	(Leitung der Dissertation)
Prof. Dr.	Josef Jiricny	(Komitee-Mitglied)
Prof. Dr.	Nicolai Bovin	(Begutachter)
Dr.	Darlene R. Goldstein	(Begutachter)

Zürich, 2010

Declaration

I declare that the present thesis was composed by my self
and the enclosed experimental work was performed by my own.
Exceptions are explicitly stated in the text.
This dissertation has not been submitted for any other degree
or professional qualification except as specified.

Francis Jacob

Zurich, 2010

TABLE OF CONTENTS

1 Introduction

The work presented in this PhD thesis is placed within translational research and focuses on biomarker detection in epithelial ovarian cancer. It combines research in cancer medicine and biology, bioorganic chemistry, glycobiology, immunology, bioinformatics and biostatistics. The first section of the introduction describes the main knowledge of epithelial ovarian cancer, including epidemiology, clinical and pathological characteristics, and summarizes the current state of biomarker research. The second part describes basic glycobiology and links aberrant glycosylation to cancer.

1.1 **Epithelial Ovarian Cancer**

1.1.1 Epidemiology and Risk Factors

From 2002 to 2006 the world wide average incidence (Figure 1) of malignant tumors of the ovary was 10.7/ 100 000 women, with a median age at diagnosis of 63 years. The risk of ovarian cancer increases steadily with age, with the highest risk occurring after menopause. Industrial countries in Europe, the USA and Israel have a high incidence of ovarian cancer with the lowest occurrence in Japan, compared to the low incidence observed in emerging and developing countries. In Switzerland epithelial ovarian cancers (EOC) are diagnosed on average in 589 women per year (1), and it is the sixth most common cancer affecting women worldwide (2). An already relative 5 year survival rate of 45.9% (Figure 1) dramatically decreases to only 20% survival in women with advanced EOC, which is present in 75% of all patients diagnosed with an EOC. Thus, EOC is associated with poor survival and represents the leading cause of death from a gynaecologic malignancy.

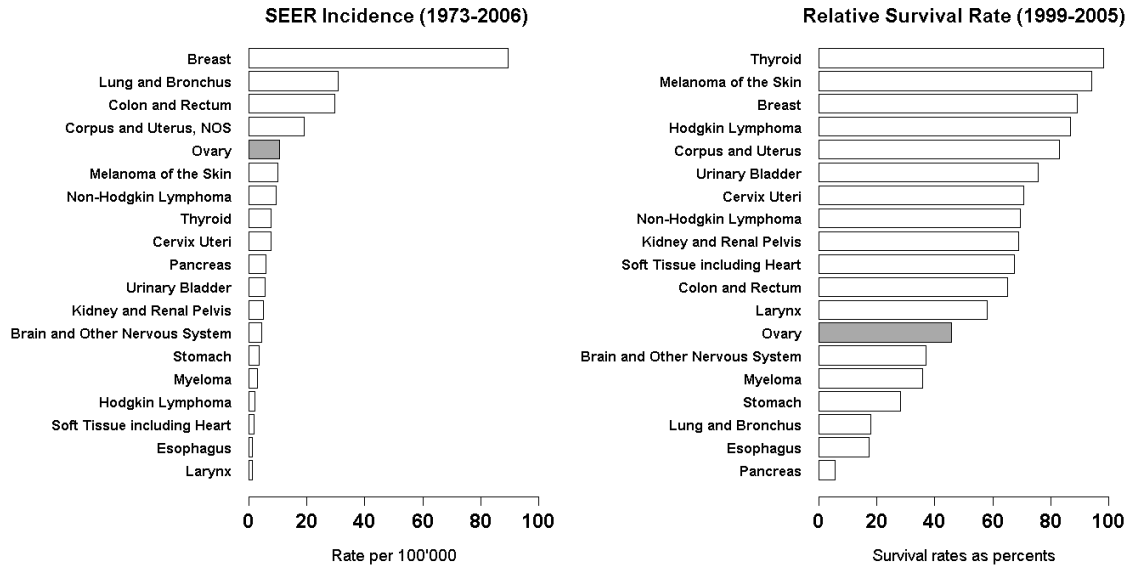


Figure 1 Incidence and relative survival rate of epithelial ovarian cancer.
Bar plots show common malignancies among women based on the Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov) database. SEER incidence between 1973 and 2006 are shown in the first bar plot. Relative survival rate as seen in the second bar plot is based on the November 2008 submission of data from the population-based SEER database including 17 registries with follow-up of patients to 2006. Malignant tumors of the ovary are highlighted in dark grey.

Several potential risk factors of ovarian cancer have been proposed based on the SEER-Stat Databases (www.seer.cancer.gov) epidemiological data. One theory suggests involvement of inflammation as a common risk factor of ovarian cancer. A recent study examined the role of talc use, history of endometriosis and use of non-steroidal anti-inflammatory drugs

(NSAIDs). It was found that the risk of ovarian cancer increased significantly with increasing frequency and duration of talc use. A history of diagnosed endometriosis was also significantly associated with a higher risk for subsequently developing ovarian cancer. Further, women who were talc users and had a history of endometriosis showed a 3-fold increased risk (3). Contrary to the initial hypothesis that risk of ovarian cancer may be reduced by use of NSAIDs, it was found that in fact the ovarian cancer risk rose with increasing frequency and years of NSAIDs application (3). Additional environmental factors such as asbestos exposure, pelvic inflammatory disease and mumps were compatible with the theory that inflammation may be a common pathway for ovarian carcinogenesis (3-6). As a path of transmission, talc and asbestos found in sanitary pads are suspected to ascend through the fallopian tube to the peritoneal cavity, which is why the preventative factor tubal ligation may reduce exposure to these and other external carcinogens.

The remaining two long-held hypotheses are incessant ovulation (7) and gonadotropin (8) theories. Both propose that continuous ovulation and gonadotropin hormones (follicle-stimulating hormone) stimulate cell proliferation resulting in malignant transformation of the ovarian surface epithelium (OSE). A protective effect of anovulatory events, such as the oral contraceptive pill, pregnancy and lactation are in concordance with these findings (9, 10). However, none of these ovulation suppression events alone is sufficient to explain development of EOC. Despite anovulation due to breastfeeding and pregnancy, similar cancer risk factors were observed in both pre- and post-menopausal women. Lifetime ovulation is suspected to be involved mainly in the pathogenesis of pre- but not post-menopausal ovarian cancer patients, whilst the protective effects of years of anovulation seem to persist till post menopause (11).

Hereditary defects are other critical risk factors. Multiple cases of breast and ovarian carcinomas within an individual family are the cardinal indicators of a possible *BRCA1* or *BRCA2* gene mutation (12). Women with these mutations inherit a 5- to 20-fold increased risk of developing breast and ovarian cancer (13). Hereditary *BRCA1* or *BRCA2* mutation-related EOC tends to occur at an earlier age than sporadic tumors, and are more often high-grade serous ovarian cancers with p53 protein dysfunction (14).

1.1.2 Screening and Management

Despite the genetic distinction between various EOC histological subtypes (15), they are always treated in the same manner. The standard initial management of EOC consists of surgical staging or debulking including total abdominal hysterectomy, bilateral salpingo-oophorectomy, infra-colic omentectomy and possible pelvic and para-aortic lympadenectomy

or appendectomy. The aim of this operation is maximal cytoreduction (16, 17) leaving no residual tumor or individual deposits less than 1cm in diameter. Optimal cytoreduction with residual volumes less than 1cm in size is associated with a markedly improved disease-specific and relapse-free survival (18).

Post-operative treatment includes six cycles of intravenous Carboplatin and Paclitaxel chemotherapy (19-21) which in some cases of optimal debulking is partially given intra-peritoneally. Since the introduction of platinum-based chemotherapies this class of drugs has shown maximal potency for ovarian cancers (19).

Nowadays, detection of EOC is usually performed using tests that examine the ovaries, pelvis, blood and ovarian tissue. A pelvic assessment includes bimanual examination of the vagina, cervix, fallopian tubes and rectum. Further, abdominal and transvaginal ultrasound examinations are performed to image the internal genital organs. Biomarker CA125 levels are usually measured to evaluate protein concentrations in blood serum and consolidate suspect findings of the clinical examinations.

Early diagnosis of EOC is difficult due to the non-specificity and vagueness of symptoms. Early phase symptoms are characterized by dyspareunia, abdominal pressure and pain. These symptoms are also present in advanced disease with additional symptoms including bloating, nausea, anorexia as well as abdominal distension mostly caused by increasing ascites.

Despite enhancements in imaging technology, gray-scale transvaginal ultrasonography remains the standard in evaluation of ovarian cancer (22) which is assessed by size, mass characteristics (cystic, solid, or both), complexity (internal septae, excrescences and papillae), and the presence or absence of abdominal or pelvic fluid (ascites or blood). Gray-scale and Doppler ultrasound examinations can help to discriminate between benign and malignant adnexal masses (23, 24). Ultrasonography and computed tomography have similar sensitivity and specificity for evaluation of adnexal masses, but ultrasonography is generally more cost-effective (25).

The current tumor marker, CA125, is applied as a biochemical detection tool. It is a glycoprotein usually elevated in patients with ovarian tumors displaying low-malignant and malignant potential. Significantly increased levels of CA125 are often obscured by several benign conditions including cirrhosis, disease involvement of a serosal surface, endometriosis, pelvic inflammatory disease, uterine leiomyoma as well as malignant cross-indications caused by breast, lung, endometrial and pancreatic cancer (26). Nevertheless, CA125 has proved to be a useful serum tumor marker for the purpose of monitoring response to chemotherapy,

detecting disease recurrence, distinguishing malignant from benign pelvic masses, and potentially improving clinical trial design (27).

1.1.3 Clinical and Pathological Characteristics of Epithelial Ovarian Tumors

Tumors of the ovary are named according to the type of cells the tumor originated from. There are three main types of ovarian tumors:

1.) stromal tumors arising from the connective tissue of the ovary,

2.) germ cell tumors derive from the gonadal tissues, and

3.) epithelial ovarian tumors are thought to originate from the ovarian surface epithelium. These tumors are the most common tumors of the ovary, representing >90% of all cases.

1.1.3.1 Grading

Based on the proliferative spectrum of epithelial ovarian tumors, and recommendations from the International Federation of Gynaecology and Obstetrics (FIGO) and World Health Organization (WHO), subgroups are categorized into benign, borderline (BL, also referred to as 'proliferating tumor' or 'tumor of low-malignant potential' (LMP)) (28) and malignant categories. Malignant tumors are classified by their histopathological grade. Grading schemes analyze the microscopic architectural and cellular characteristics, such as tubule formation, mitotic rate and nuclear pleomorphism. However, due to lack of standardization, applying one single grading system to all the histological subtypes of EOC is difficult and clear cell ovarian cancers are therefore not graded at all. EOC with well differentiated cells are classified as the lowest grade (grade 1). Grade 2 describes moderately differentiated cells and in highest grades (grade 3 and 4) tumor cells are poorly differentiated or undifferentiated. In the majority of these cases, a clear separation is difficult and thus grades 3 and 4 are usually combined. Higher grades are associated with an increased malignant potential of cancer cells, as well as the ability to spread faster and an increasing potential for disease recurrence. The chance of a better outcome or successful treatment is ameliorated if the grade of EOC is lower.

1.1.3.2 Histological Subtypes and Their Pathological Morphology

The heterogeneity of epithelial ovarian tumors is exhibited by the appearance of diverse histological subtypes classified according to their patterns of differentiation (27, 29, 30). The following paragraphs briefly highlight the common histotypes of ovarian tumors and their pathological morphology with the main focus on EOC.

1.1.3.2.1 Serous Ovarian Tumors

Serous tumors of the ovary are usually classified based on the proclivity of their epithelium to produce fine papillary structures. Their epithelial cells resemble those of the native fallopian tube mucosa (Figure 2 c). Cells of moderately and poorly differentiated serous ovarian cancers (SOC) (Figure 2 b) are small with round to slightly oval nuclei and prominent nucleoli. The cytoplasm is generally scant and basophilic or amphophilic with indistinct cell borders (31). A characteristic of SOC is the presence of randomly distributed single mononuclear tumor giant cells, sometimes clustered in areas of ischemic degeneration. High protein expression of cytokeratins, epithelial membrane antigen (EMA) and milk fat globule-EGF factor 8 protein (HMFG-IIIC12) are immunohistochemical features found in SOC cells (31). SOC accounts for approximately 60% to 80% of all EOC and is by far the most aggressive histological type (32). Recent findings distinguish low and high grade SOC subtypes. High grade SOC involves the surface of the ovary, often bilaterally, along with the peritoneal membranes, with rapid onset of carcinomatosis, a fact that restricts the surgical option for complete debulking (32). Low grade SOC often originates from a serous borderline tumor. These tumors can be reproducibly differentiated from high grade tumors, based primarily on their very uniform nuclei, using low mitotic rate as a secondary diagnostic feature (33, 34).

1.1.3.2.2 Endometrioid Ovarian Tumors

Endometrioid tumors of the ovary are very rarely found in benign and low-malignant conditions. It is the second most common subtype of EOC representing approximately 20% of all cases (Figure 2 d). The majority of reported cases of endometrioid histotype are well or moderately differentiated endometrioid EOC (EnOC) (Figure 2 b). Low grade EnOC are strongly associated with endometriosis. Similar molecular genetic alterations, including loss of heterozygosity at 19q23 and mutations in *PTEN*, have been reported in endometriosis, atypical endometriosis and EnOC in the same specimen (35-38). EnOC are macroscopically characterized by cysts with a smooth outer surface and soft masses or papillae partly filling cystic spaces. EnOC display histological variability showing microglandular or trabecular patterns and elongated nuclei with coarse granular chromatin. Their morphological patterns tend to resemble those of proliferative endometrium (Figure 2 c). Using immunohistochemical techniques, EnOC also stain positive for cytokeratins and epithelial membrane antigen (31).

1.1.3.2.3 Mucinous Ovarian Tumors

Benign mucinous tumors are mostly cystadenomas and are characterized by a uniform single layer of tall columnar cells with pale, amphophilic cloudy cytoplasm and small basal nuclei. Mucinous ovarian cancers (MOC) are much less common than was previously assumed, as historically, many case series included specimen of metastatic carcinoma with mucinous differentiation which were found to be primary gastrointestinal or hepato-biliary cancers (39). The morphology of benign as well as (low-) malignant mucinous cells resemble those of the endocervix and gastrointestinal mucosa (31) (Figure 2 b, c). The large size of MOC and their unilateral presentation underscores the fact that they grow slowly (40). It has been suggested that MOC develop slowly, in a stepwise fashion from benign precursors. This was emphasized by molecular genetic studies which have shown that the most common molecular genetic alterations in mucinous BL and MOC is a point mutation in *KRAS* (41-43). MOC are estimated to make up maximally 10% of all EOC. Malignant tumors of mucinous subtype are predominantly multilocular cystic neoplasms. They consist of solid areas and luminal nodules which are more common than in the less proliferative lesions, as haemorrhage and necrosis are (31). Microscopically, MOC show proliferating and benign areas. Invasive areas usually resemble mucin-secreting adenocarcinomas of intestinal origin. Their architecture varies with the amount of mucin produced, cellular differentiation and prominence of stroma (31).

1.1.3.2.4 Clear Cell Ovarian Tumors

Clear cell tumors share the same proliferative characteristics as other epithelial ovarian tumor histotypes do. However, benign tumors and proliferating clear cell tumors are extremely rare (2-5%) (Figure 2 d). Clear cell ovarian cancers (ClCOC) are macroscopically characterized by large thick-walled unilocular cysts containing blood-stained fluid with polyploidy nodules. They usually consist of small to large layers of polyhedral clear cells separated by delicate fibro- vascular septa. Their cells have distinct cell borders and eccentric rounded or slightly angular nuclei (31). ClCOC are also strongly associated with the presence of endometriosis (39). Using immunohistochemistry, high grade ClCOC diffusely stains for cytokeratins, the CA125 antigen, as well as for vimentin.

1.1.3.2.5 Less Common Histotypes of Epithelial Ovarian Tumors

Transitional cell tumors seem to derive directly from the OSE which undergoes metaplasia to form uroethelial-like epithelium. They share cytological as well as immunohistochemical similarities with the epithelium lining the urinary tract (44). It resembles those of the Walthard rests, which are epithelial inclusions found most commonly beneath the

serosal surface of the fallopian tubes. Transitional cell ovarian cancers (TCOC), also referred to as malignant Brenner tumors, show a heterogeneous solid malignant proliferation with variable amounts of intervening stroma.

Mixed epithelial tumors are defined by the presence of two or more of the epithelial elements described above. They are found throughout the full proliferative spectrum from benign tumors to invasive carcinomas. A mix of serous and endometrioid characteristics is most commonly observed in benign tumors, whereas malignant tumors with mixed histotypes consists of SOC/ EnOC and EnOC/ ClCOC (31).

Undifferentiated ovarian carcinomas (Figure 2 d) include tumors with minimal differentiation such as occasional endometrial glands, psammoma bodies and droplets or pools of mucin. Macroscopically, tumors tend to be large and solid with haemorrhage and necrosis commonly present. Several histopathological features are detectable with epithelial cell markers such as cytokeratins staining positively in following immunohistochemical examination (31).

1.1.3.3 Heterogeneity and Determination of Morphological Identity

A paradigm shift has happened in EOC during the last decade in that this formerly uniform disease was recognised to be of heterogeneous nature. Classification of these already heterogeneous tumors is further complicated by the presence of fallopian tube and primary peritoneal cancers. In some cases, at advanced stages, it is difficult to determine if the cancer originated from the ovary, tube or peritoneum. Their morphological and clinical similarity to EOC is possibly due to the OSE, the peritoneal, and the fallopian tube epithelia all sharing the same embryonic precursor (Figure 2 a, c). Observations suggest that some EOC originate from the fimbria of the distal tubes (45). OSE and the Mullerian ducts, a tissue of mesodermal origin that develops to form the fallopian tubes, uterus, and the upper portion of the vagina, are derived from the embryonic coelomic epithelium (46, 47). OSE is a simple cell surface monolayer that resembles the mesothelial lining of the abdominal cavity, whereas Mullerian epithelium is organized in tubal or glandular structures. The development of the reproductive tract is tightly regulated by *homeobox* genes (48). Many morphologic characteristics present in EOC are similar to those of the Mullerian lineages, *e.g.* EnOC are characterized by endometrial-like structures, SOC resemble malignant tumors of the fallopian tube (Figure 2 b, c), and MOC are composed of intestinal- or endocervical-like cells (49, 50) (Figure 2 b, c). This theory was confirmed by gene expression analysis on *homeobox* genes (Table 1). Genes that are exclusively expressed in Mullerian duct cells were also found in EOC subtypes according to

the pattern of Mullerian-like differentiation, but were not expressed in OSE (51, 52) (Table 1). The lack of an explanation for the Mullerian-like features of EOC has stimulated the debate challenging the traditional notion that the OSE is the cell-of-origin of these cancer types (53, 54).

***homeobox* gene**	**usually expressed in ***	**EOC expression**	**stable expression in transformed mouse OSE induces formation of**
HOXA9	fallopian tube, endometrium, endocervix	SOC, EnOC, MOC	cystic, papillary tumors resembling SOC
HOXA10	endometrium, endocervix	MOC, EnOC	glandular tumors resembling EnOC
HOXA11	endocervix, lower uterine segment	MOC	tumors resembling MOC

Table 1 *HOX* gene expression in epithelial ovarian cancer.
*Weak or no staining for homeobox proteins is present in healthy OSE cells. Abbreviations: serous (SOC), endometrioid (EnOC), and mucinous (MOC) ovarian cancer; Table is adapted from Hennessy *et al.* 2009 (18).

Several mouse genetic models have supported OSE as the origin of EOC (55, 56). Auersperg *et al.* 2001 have speculated that Mullerian metaplasia is possibly connected to ovarian neoplasia, and that OSE cells gain a growth advantage by acquiring Mullerian-like phenotypes. They have speculated that Mullerian-like differentiation of OSE cells may increase hormone responses, because estrogens have mitogenic effects on Mullerian epithelium but not on OSE. Unlike Mullerian epithelium, OSE only has a fragile attachment to underlying stromal components, and therefore Mullerian-like differentiation may enhance epithelial–stromal interactions. These explanations for Mullerian-like involvement in differentiation of OSE cells is interesting but has yet not been investigated further.

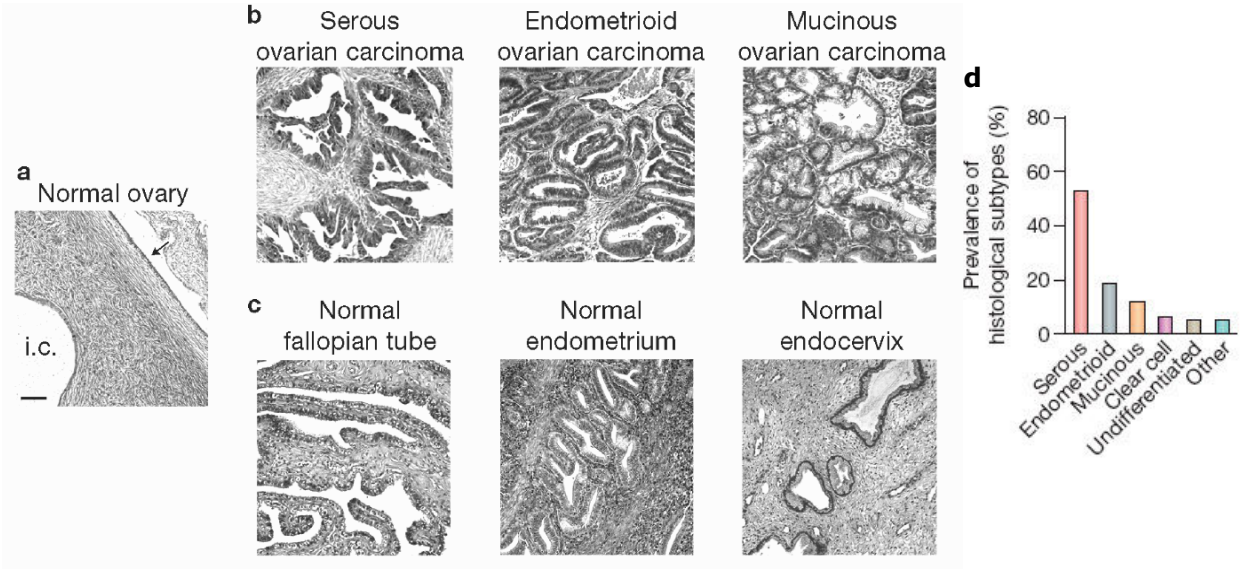


Figure 2 Histopathology and morphological features of common epithelial ovarian cancer subtypes. Hematoxylin–eosin-stained sections of (a) normal human ovary surrounded by a monolayer epithelium (arrow) and containing an inclusion cyst (i.c.), (b) sections of ovarian carcinomas: serous, with papillary features; endometrioid, with glandular features; and mucinous. (c) Tissue sections of normal fallopian tube epithelium, endometrium and endocervix. (d) Distribution of histological subtypes. Figure adapted from Naora *et al.* 2007 (50). Black bar (a) indicates 50 µm.

1.1.3.4 Staging

The FIGO staging system has been structured to uniformly represent the major prognostic factors in predicting patient outcome and lending order to the complex dynamic behaviours of gynaecologic cancers. The FIGO system offers a classification of gynaecologic cancers in a unique setting and thus allows comparisons between clinical cases (57). Proliferating and invasive malignant tumors of the ovary are staged (FIGO) as: Stage I A-C, Stage II A-C, Stage III A-C, and Stage IV. Briefly, Stage I tumors are confined to one or both ovaries. Tumors that have spread to pelvic organs (*e.g.* fallopian tubes, uterus) without metastasis to abdominal regions are defined as Stage II. At Stage III, one or both ovaries are affected by malignant tissue with additionally at least microscopically confirmed peritoneal metastasis outside the pelvis and/ or regional lymph node metastasis. Stage IV tumors of the ovary have metastases in other organs outside the peritoneal cavity.

1.1.4 Genetic Profiling Leads to Development of New Epithelial Ovarian Cancer Models

Based on the strong heterogeneity of EOC and its unknown cell of origin, it is not surprising that new hypotheses are continually arising. Due to the human genome project and the era of high-throughput profiling technologies we now have indications that certain histological subtypes have a genetically distinct profile (58). Based on recent findings, this

section describes the main manifested theories regarding epithelial ovarian cancer development.

One of the most important recent observations is that low grade serous carcinoma (LGSC) and high grade serous carcinoma (HGSC) are different tumor types (59, 60). In 70% of all investigated cases serous BL tumors and LGSC were characterized by specific mutations in *BRAF* and *KRAS*. Additional patterns of *p53* mutations separate serous BL tumors and LGSC, compared to HGSC and support a new model of ovarian carcinogenesis (40, 60, 61). *P53* gene mutations were shown to be associated with a serous BL components and do not seem to be linked to HGSC (62-65).

Experimental findings have led to the creation of a new model for classifying EOC. In this model, two main subclasses have been proposed- Type I and Type II tumors. This categorization comprises distinct patterns of tumor progression and molecular genetic changes (61, 66). Type I tumors include LGSC, low grade EnOC, MOC, and a subset of ClCOC, which evolve in a stepwise fashion from well recognized precursors, in most cases, BL tumors. They are mostly characterized by slow growth. The latter, in turn, appear to develop from the OSE or inclusion cysts (Figure 2 a) as is the case for SOC and MOC. In this model endometriosis is an associated precursor of EnOC and ClCOC (35-39). In contrast, Type II tumors are high grade and have generally spread beyond the ovaries at presentation. HGSC, high-grade EnOC, undifferentiated EOC and malignant mixed mesodermal tumors are considered to comprise Type II group tumors. It is thought that Type II tumors are rarely associated with morphologically recognizable precursor lesions and may arise from “dysplasia” in inclusion cysts or serous intraepithelial carcinoma in the fallopian tube (67). Type I is usually associated with mutations in *BRAF*, *PIK3CA*, *ERBB2*, *KRAS*, *CTNNB1* and *PTEN*. Type II tumors are generally characterized by a high frequency of *p53* mutations which occur rarely in Type I tumors. This definition is consistent with a recent publication, but indicates that changes between different histological subtypes are associated with specific gene expression patterns regardless of the substage and grade of the disease (68).

Another model focuses primarily on MOC, which appears to have weaker associations with reproductive factors (69), a positive relationship with smoking that is not observed for other histological subtypes (70), and no association with *BRCA* mutations (71). Their clinical and morphological distinction was described and compared to other histological types of EOC. Initial gene expression profiling using oligonucleotide arrays and validation by real time polymerase chain reaction (PCR) revealed a distinct increase in gene expression of galectin 4 in MOC compared to other EOC histotypes comprising SOC, EnOC and ClCOC (72). Recent

findings based on immunohistochemical analysis support these observations. Zinc finger-containing GATA transcription factors GATA4 and GATA6 were found to be lost (80%-90%) in all histological subtypes except MOC (11/12 cases), where the presence of both transcription factors were observed (73). Gene expression patterns of early stage EOC also revealed a clear distinction between MOC and other EOC histological subtypes (68).

1.1.5 Ovarian Cancer Biomarker Detection

One of the most promising aims to manage EOC is early detection. Early stage EOC can be cured with at present available cytoreductive surgery and adjuvant chemotherapy in more than 90% of treated patients. So far, CA125 is currently the best biomarker for early detection. It is a heavily glycosylated mucin (MUC16) recognized by a murine monoclonal antibody (mAb) OC125 (74). This mAb is used in a blood serum-based immunoassay to detect elevated CA125 protein levels preferentially in non-mucinous EOC patients (75). To increase sensitivity, an advanced second-generation CA125 II sandwich Enzyme-Linked Immunosorbent Assay (ELISA) has been widely accepted. It utilizes a distinct mAb to capture the protein followed by the use of the original OC125 antibody for detection (76), which improves measurement characteristics at low CA125 levels.

During recent years, enormous effort has been made in the field of early detection of EOC. Due to the lack of sensitivity and specificity, CA125 as a sole biochemical biomarker is not recommended to screen asymptomatic women for EOC (77). Therefore, novel biochemical candidates alone and in combinations which have been previously reported are mentioned and described in this chapter.

1.1.5.1 Single Biomarker Detection

During the last 20 years, a large number of serum tumor markers have been evaluated for their ability to detect EOC at an early stage (Table 2). They were discovered through multiple approaches, cDNA and oligonucleotide arrays have identified a number of novel biomarkers, including *HE4* (78, 79), *prostasin* (80) and *osteopontin* (81) (Table 2).

Another study used a cDNA database mining strategy followed by real time quantitative reverse transcription PCR, which identified *VTCN1* (*B7-H4*) (Table 2) as a novel gene overexpressed in EOC compared to normal tissue samples (82). Immunohistochemistry revealed concordant findings and observed a predominant plasma membrane staining of B7-H4. Thus, sandwich ELISA was developed by applying two mAbs to two different B7-H4 protein epitopes (82). Ovarian tissue lysates from 251 women with EOC, benign tumors of the ovary, and normal ovaries were investigated and it was found that B7-H4 expression was low

in normal and benign ovaries as opposed to high in malignant tumors, with compatible sensitivity and specificity compared to CA125 (83).

Another strategy to identify biomarkers used hybridoma technique to generate mAbs that bind to antigens which were found to be present on the surface of mesothelial mesothelioma and ovarian cancer cells (84). Mesothelin, a glycosyl-phosphatidylinositol-anchored glycoprotein present on the cell surface, was discovered *via* this method. Detection of its soluble form using receiver operating characteristics (ROC) analysis showed similar detection ability to that of CA125. Improved detection of EOC was achieved using a combination of both mesothelin and CA125 (85). Recent findings support an association of mesothelin expression to poor survival when examined by real-time quantitative reverse transcription PCR (86) and ELISA (87).

A laboratory investigation into the biology of the epidermal growth factor receptor (EGFR) found that decreased expression of a soluble fragment of EGFR is present in serum samples obtained from EOC patients (88). Measured concentrations of its 110kDa fragment were found to be more effective in discerning advanced from early stage EOC, as well as from healthy women, however, no comparison to CA125 was made (89).

During the last 25 years, despite the intensive screening for other single promising biomarkers of EOC, CA125 still holds the leading position in this field. Single biomarker candidates are selected by their ability to detect EOC with comparable sensitivity and specificity to CA125 (Table 2). Novel single biomarkers and CA125 have been combined by different statistical algorithms and reveal a slightly better discrimination than either marker alone (85, 90-92). Biomarker validation was predominantly performed on biological fluids, *e.g.* serum, plasma, and urine. To date sandwich ELISA is the leading technique in single biomarker measurement, which can be explained by its antigen accessibility. It is also less time consuming and more affordable. Using sandwich ELISA, studies have measured the amount of antigen between two layers of antibodies; mostly two mAbs directed to different antigenic sites of a single biomarker protein. Interestingly, all ten highlighted candidates consist of multiple post-translational modifications, however only glycosylation is present in all candidates (Table 2).

Ref.	Protein name	Gene name	PTM	Molecular function
(75)	CA125	*MUC16*	disulfide bond, glycoprotein, phosphoprotein	protein binding
(90, 91)	HE4*	*WAP5, WFDC2*	disulfide bond, glycoprotein	serine-type endopeptidase inhibitor activity
(82, 83)	B7-H4*	*VTCN1*	disulfide bond, glycoprotein, lipoprotein	negatively regulates T-cell-mediated immune response
(93)	(total) Inhibin*	*INHA & INHB(A-C)*	disulfide bond, glycoprotein, lipoprotein	growth factor, hormone, cytokine, signal transducer
(94)	Kallikrein8	*KLK8*	disulfide bond, glycoprotein, zymogen	protein binding, serine-type endopeptidase activity
(95)	Mesothelin^{+}	*MSLN*	cleavage on pair of basic residues, GPI-anchor, glycoprotein, lipoprotein	protein binding
(96)	Nidogen-2*	*NID2*	disulfide bond, glycoprotein	calcium ion binding, collagen binding
(81)	Osteopontin^{+}	*SPP1, BNSP, OPN*	glycoprotein, phosphoprotein	cytokine activity
(80)	Prostasin*	*PRSS8*	disulfide bond, glycoprotein, zymogen	serine-type endopeptidase activity
(88, 89)	c-ErbB-1	*EGFR, ERBB1*	disulfide bond, glycoprotein, isopeptide bond, phosphoprotein, ubiquitin conjugation	broad spectrum

Table 2 Potential biomarkers for the detection of epithelial ovarian cancer.
Biomarkers that compete with CA125 for early detection of EOC are shown. Despite different ways of biomarker identification, all listed proteins were measured by sandwich ELISA. All proteins contain glycosylation as their main post-translational modifications (PTM). Protein names, gene names, PTM and molecular function were taken from UniProt (http://www.uniprot.org/). Reported to complement CA125 in the same study* or in a later study^{+}.

1.1.5.2 Strategies to Improve Biomarker Detection

Novel technologies have been developed to detect new biomarkers for early detection of EOC. In a recent publication, concentrations of leptin, prolactin, osteopontin, insulin-like growth factor II, macrophage inhibitory factor, and CA125 were examined using a multiplex, bead-based immunoassay system. All six biomarkers exhibited the same pattern for both the ELISA and multiplex system. Several models using biomarker combinations revealed a higher sensitivity (84%-98%, specificity 95%) than CA125 alone (sensitivity 72%, specificity 95%) (97). Another study simultaneously investigated 204 molecules in 294 serum samples by multiplex, bead-based immunoassay. A panel of putative biomarkers was found, including CA125 which showed the best discriminative power (98).

To improve the diagnostic performance of single biomarkers, new technologies based on high-throughput profiling have emerged during recent years. Serum profiling is mostly

based on proteomics (58), the study of an entire set of proteins expressed by its genome. Proteomic approaches display a broad spectrum of technical and biological variables. Mass spectrometry (MS) is one of the main technologies employed in this field. Surface-enhanced laser desorption and ionization (SELDI) for detection of novel cluster patterns of low-molecular-weight moieties in serum samples from EOC patients has been reported in ROC to yield 100% sensitivity and 95% specificity (99), unfortunately no comparison to CA125 detection was made for final conclusions. Studies on pre-selected peptides employing different protein chips (*e.g.* IMAC3, WCX2 or SAX2) and/ or additional experimental steps are used to investigate a sub proteome with defined physiological or chemical properties. For example, a publication using a protein chip (IMAC3) with affinity selection for metal-binding proteins coupled to SELDI MS demonstrated that a combination of four peptides complement CA125 performance (100). Another MS approach revealed a putative tumor marker, an 11.7kDa serum protein amyloid A1 in pre-selected thermo stable plasma fractions (101). A phosphorylated fibrinogen-α-chain isoform was identified as a potential biomarker by 2-dimensional polyacrylamide gel electrophoresis (2D-PAGE) and phosphoprotein specific staining (102). Another study identified 14 putative tumor markers in serum, pre-processed by biomarker enrichment reagent to capture high-abundant carrier proteins in blood (103). Despite promising results, MS approaches did not perform better than CA125 in the identification of biomarkers.

Serological analysis of recombinant cDNA expression (SEREX) is another auspicious method to profile EOC biomarkers. To identify ovarian tumor antigens a papillary serous carcinoma cDNA library isolated from ascites was transferred into *Escherichia coli via* a bacteriophage expression system. Nitrocellulose transferred plaques were screened against a serum pool consisting of serum from 20 patients with advanced EOC. After stripping and re-probing with pooled normal sera, true-positive immunoreactive clones were identified by sequence alignment. SEREX data indicated that interleukin 8 (IL-8) could be a potential target for the host immune response, resulting in the generation of anti-IL-8 antibodies. An immunofluorescent bead-based assay was subsequently developed to detect IL-8 and corresponding anti-IL-8 antibodies. ROC analysis revealed acceptable detection performance but did not reach the same levels as those for CA125. Combining IL-8, anti-IL-8 IgG and CA125 resulted in an increased classification power compared to markers analyzed individually (104).

High-throughput profiling technologies using serum antibodies are also tools for biomarker discovery with good prospects. Recent findings demonstrated the accessibility of serum circulating antibodies to tumor-associated antigens in EOC. This is based on recent

knowledge of circulating autoantibodies related to EOC-associated antigens, *e.g.* p53, HOXA7, HOXB7, cathepsin D, MUC1, IL-8 and S100A7. These proteins are potential screening tools because they are secreted by tumor cells and are detectable by autoantibodies even when the tumor is very small and antigen expression is low (105, 106). In EOC it has been shown that p53 specific autoantibodies are present and associated with high grade disease (107). These findings support the theory of human immune recognition to known and unknown tumor-associated antigens of any type of biomolecule.

In 2006, An *et al.* published the first study, which investigated the role of the glycome in EOC, profiling the entire complement of glycans (108). Firstly, oligosaccharides covalently linked to human serum proteins and human ovarian cancer cell lines were released chemically. Changes in glycosylation were then monitored by matrix-assisted laser desorption/ ionization Fourier transform ion cyclotron resonance mass spectrometry (MALDI-FTICR-MS). Mass spectrometric analysis of purified oligosaccharides identified 15 unique serum glycan markers in all patients, but not in any of the normal controls. This publication was the first EOC biomarker detection study using a high-throughput setup to profile changes in glycosylation.

1.2 **Glycobiology and Cancer**

1.2.1 Glycosylation as the Main Post-Translational Modification Event

The term “Glycobiology” was first coined in 1988 by Rademacher, Parekh and Dwek and is defined in its broadest sense as the study of the structures, biosynthesis, and the biology of saccharides (sugar chains or glycans) which are widely distributed in nature. In eukaryotic cells, the main mechanisms by which cell surface associated components become glycan-coated occur in the endoplasmic reticulum–Golgi biosynthetic pathway (109). Glycans on cell surfaces, usually named ‘glycocalyx’ (Figure 3 A) were first mentioned by a electron microscopists (110). All forms of molecular modifications by oligosaccharides have the inherent potential to add new information to the carrier molecule which usually comprises a lipid or protein backbone. More than half of all proteins are glycosylated as estimated from the SWISS-PROT database (111). Oligosaccharide trees are not primary gene products, but are instead built by catalytic reactions of glycan processing enzymes, glycosidases and glycosyltransferases, in complex with high energy sugar nucleotide donors. Human cells possess a repertoire of about 250 glycan related enzymes, and the co-ordinated action of as many as thirty enzymes may be involved in the synthesis of a single complex oligosaccharide (112).

The most common monosaccharide units of animal glycoconjugates (=carbohydrate covalently linked to other chemical compounds) appear naturally linked to proteins as *N*- or *O*-linked saccharides, or to lipids *via* a glycan bridge between phosphatidylinositol and a phosphoethanolamine named glycosphingolipids. An *N*-linked oligosaccharide is a sugar chain covalently linked *via N-acetyl glucosamine* (*GlcNAc*) to an asparagine residue of a folded polypeptide chain within the consensus sequence, Asn-X-Ser/Thr. The synthesis of such an oligosaccharide is performed on the cytoplasmic side of the endoplasmic reticulum (109). An *O*-glycan is typically linked to a polypeptide with motif Ser/Thr, *via N-acetyl galactosamine* (*GalNAc*). There is no defined consensus sequence regarding *O*-linked glycosylation, however, glycosylated Ser/Thr residues are often located in Prolin-rich sequences. *O*-linked glycosylation of glycoproteins is based on post-translational modifications (PTM) which begin in the *cis*-compartment of the Golgi apparatus. Other glycopeptide linkages have been described, and thirteen different monosaccharides and eight amino acids are known to be involved in glycoprotein linkages among all organisms. If the anomeric configurations, the phosphoglycosyl linkages, as well as the GPI (glycophosphatidylinositol)

phosphoethanolamine bridge are also considered, the observation results in a total of at 41 possible bonds (113).

The genome encoding deoxyribonucleic acid (DNA) is a double-stranded, helical polymer of nucleotides (=phosphate group + deoxyribose sugar + heterocyclic nucleotide base) that is held together by two or three hydrogen bonds between the nitrogenous bases, adenine-thymine and cytosine–guanine, respectively. In comparison, a protein (=entire set of proteins is named proteome) is a linear polymer of twenty common amino acids linked *via* a peptide bond. Additional biochemical information in proteins lies in various intra- and inter-molecular bonds (*e.g.* disulfide bridges, hydrogen bonds) and in the numerous post-translational modifications (PTM). All these interactions and modifications, together with the final tertiary structures of a protein, dramatically enhance their functional diversity (Figure 3 B). Among the PTM, protein and lipid glycosylation is the major class, and the glycans represented in the glycome (=entire repertoire of glycans and glycoconjugates of the organism) provide biological access to vast amounts of information at minimum genetic cost (114). Various branched or linear glycosylation trees of yet unknown length are assembled from ten monosaccharides: sialic or *neuraminic acid* (*Neu5Ac*), hexoses *glucose* (*Glc*), *galactose* (*Gal*), *fucose* (*Fuc)* and *mannose* (*Man*), hexosamines *GlcNAc* and *GalNAc*, and less commonly *xylose* (*Xyl*), *glucuronic acid* (*GlcA*) and *iduronic acid* (*IdoA*) (Figure 3 C). Generally, most of these monosaccharides are linked *via* alpha (*α*) or beta (*β*), *1-2*, *1-3*, *1-4* or *1-6* glycosidic linkages. In conclusion, glycosylation is therefore more abundant and structurally diverse than all other types of PTM combined (115, 116). Based on the potential chemical information content of the glycome (Figure 3) due to its structure and biological involvement, it is not surprising that glycosylation has been linked to a variety of cancers.

Glycosylation is a common biochemical modification of proteins and lipids, it is important for stability, solubility, secretion of signals, regulation of interactions, extracellular recognition, and folding (117). Glycan involvement in host-pathogen or -virus interactions occurs over a broad biological spectrum. For example, adaptation of hemagglutinin binding of avian influenza A viruses to *α2-6* sialylated oligosaccharides located in the upper respiratory epithelia (118); and glycosaminoglycans and proteoglycans presented on host cell surfaces as substrates for the attachment of bacterial pathogens (119, 120). Interestingly, in this context, Lewis blood group antigens which are usually located in human gastric mucosa are also present on *Helicobacter pylori* lipopolysaccharide structures to camouflage the bacterium and facilitate their initial colonization (121). Glycans have also been described in inherited diseases, *e.g.* Tn (*GalNAcα*) polyagglutinin syndrome (122) or Wiskott-Aldrich syndrome which are

characterized by markedly increased branching of *O*-glycans on sialophorin, also known as leukosialin or CD43, a major glycoprotein of lymphocytes (123).

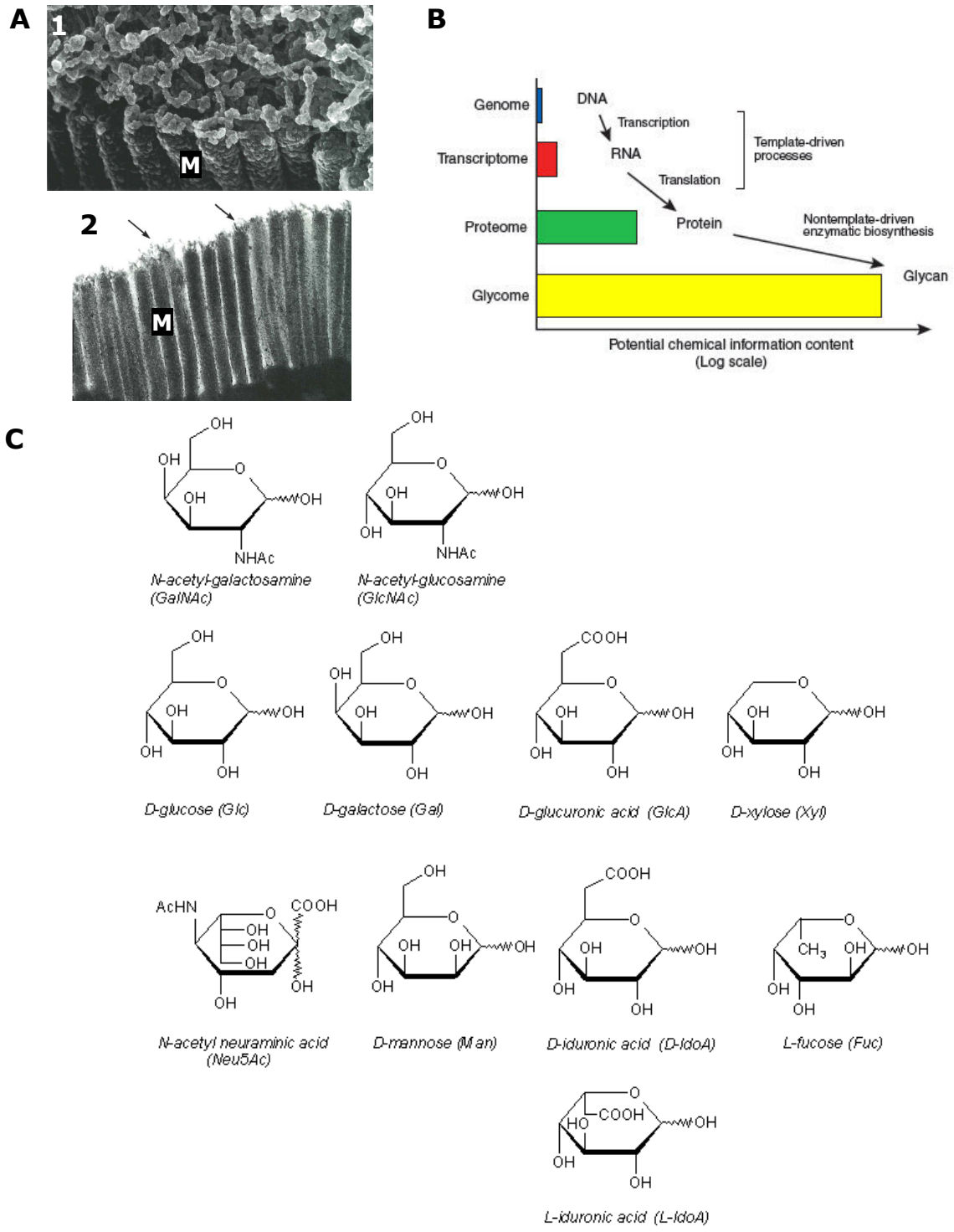


Figure 3 Relative information content of the glycome.
Scanning electron microscopy of the villous surface of the mouse small intestine shows the 'glycocalyx' (A). A1- layer of enteric surface glycosylation (x 200 000), A2- glycan surface coat on the microvilli (M) of the villous epithelial cells (x 77 000) (124). Glycome enhancement of the molecular and functional diversity of the proteome on a log scale of potential chemical information content (114) (B). Glycan structures consist of common monosaccharides which appear in higher mammals (C).

1.2.2 Tumor-Associated Carbohydrate Antigens

Glycosylation is a universal feature of all cancer cells, but only certain distinct glycan alterations are frequently associated with oncogenic transformation. Aberrant glycosylation occurs in essentially all types of human cancers, and many glycosyl epitopes constitute tumor-associated carbohydrate antigens (TACA) (125).

1.2.2.1 *O*-linked Core Structures

To date, there are only a few known cancer-related glycans including the 'core' *O*-glycan family structure Tn (*GalNAcα*) antigen, one of the most specific human cancer-associated structures. 'Core' structures are short *O*-linked glycans with the largest being trisaccharides (Figure 4). Frequently exposed Tn is caused by increased gene expression of pp*GalNAc*-transferase that attaches the first *GalNAc* to a polypeptide. High Tn antigen expression also correlates with decreased activity of *β1-3* galactosyltransferases, catalyzing the common core 1 *O*-glycan structure *Galβ1-3GalNAc*-Ser/Thr. The core 1 is the precursor for mucin-type *O*-glycan structures in animal cell surfaces and secreted glycoproteins (109).

Association of breast and colon cancer with sialylated Tn antigens (sTn/ *Neu5Acα2-6GalNAc*) (Figure 4) was described as a result of observed competition between *α2-6* sialyltransferases and *β1-3* galactosyltransferases (126, 127). Sialylated Tn immunohistochemical expression is highly variable in breast cancer. However, it is commonly accepted that at least 25-30% of breast tumors are sTn antigen positive. Its expression has also been proposed to be an independent predictor of poor prognosis in breast cancer patients (128).

The Thomsen-Friedenreich antigen (core 1, T antigen, TF/ *Galβ1-3GalNAc*) is another glycan belonging to the 'core' structures (Figure 4). This antigen is usually only an intermediate structure which is catalyzed to a core 2 (*GlcNAcβ1-6(Galβ1-3)GalNAc*). Due to its aberrant glycosylation, TF has been described as a common glycan of malignant transformation with strong immunohistochemical expression in several cancer cell membranes (129). In contrast, normal cell surfaces revealed discrete expression of TF. Another study performed on colon cancers found that increased expression of TF and core 2 is associated with decreased core 3 expression (130). MS and nuclear magnetic resonance analysis independently confirmed that core 3 structures are expressed in mucins of the descending normal colon (131).

Other 'core' structures are less common, like core 5 antigen which was described in a rectal adenocarcinoma mucin-type glycoprotein (132) and in human meconium (133). Controversially, a recent study found core 5 glycan structures in mucins from different parts of the normal human intestinal tract. Therefore, they suggested that core 5 can not be used to

diagnose a pathological case (134). Core 6 *O*-glycan formation was reported in human embryonic gut and mucins in ovarian cysts (109).

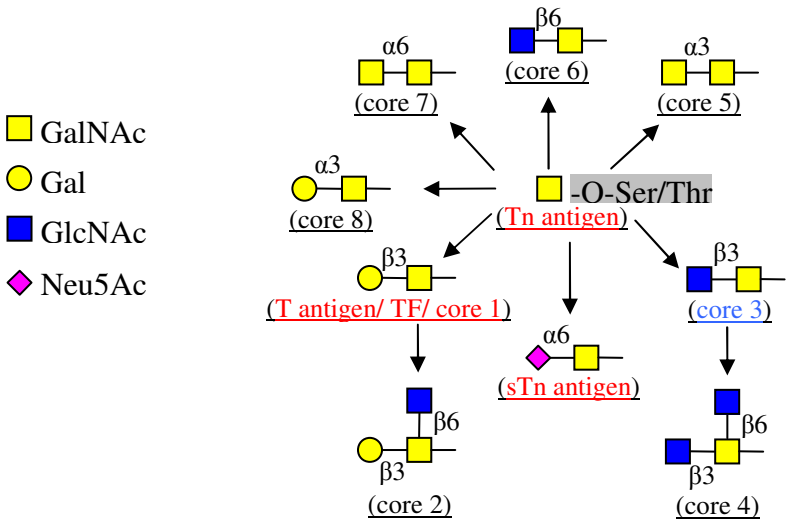


Figure 4 Pathway of *O*-linked 'core' structures.
Increasing length of *O*-linked 'core' structures descended from Tn. Additional monosaccharides are added in a stepwise manner to build glycan chains of various length and complexity. Cancer association (red) was described on Tn, sTn and TF. Normal cells (blue) exclusively express core 3.

1.2.2.2 Lewis Blood Group Antigens

The term Lewis refers to the family name of individuals who suffered from a red blood cell incompatibility that helped lead to the discovery of this frequently called "histo-blood group antigens" (109). Lewis antigens are widely distributed in human fluids and tissues (135). Expression of Lewis antigens depends on alleles inherited on two independent loci, *FUT3* (Lewis; *LE* gene) and *FUT2* (Secretor *SE* gene). Both alleles encode different fucosyltransferases that interact to form Lewis type 1 and 2. The family of Lewis blood group antigens consists of ten naturally occurring glycans with a similar structure which is based on a *α1-3(4)* linked *fucose* (Figure 5). Lewis structures are receptors for pathogenic bacteria, *e.g. Helicobacter pylori* (136), *Candida spp.* (137) and uropathogenic *Escherichia coli* (138). Their clinical significance has been described in cardiovascular disease (139), leukocyte adhesion deficiency syndrome (140), and renal transplantation (141).

An association of type 1 (*α1-3* fucosylated glycan) Lewis structure and cancer has been described. Increased expression of Lewis x (Le^x/ *Galβ1-4(Fucα1-3)GlcNAc*) was found in immunohistochemical analysis of colorectal cancers (142), renal cancers (143), and transitional carcinoma of the bladder (144). Its sialylated form, sialylated Le^x (sLe^x/ *Neu5Acα2-3Galβ1-4(Fucα1-3)GlcNAc*) was identified as a prostate specific antigen, and is a serum marker for prostate cancer (145). The presence of sLe^x was described in poorly differentiated non-small

cell pulmonary cancer and in primary liver cancer with significantly higher levels of sLe^x expression compared to low grade cancers (146).

The malignant transformation of sLe^x was further investigated by examination of a related carbohydrate determinant belonging to 6-sulfation (6' sulfo sLe^x). Glycan 6-sulfo sLe^x was preferentially expressed in nonmalignant colonic epithelial cells surrounding cancer cells, which is in contrast to non-sulfated sLe^x, expressed in cancer cells. A possible function of sulfated carbohydrate determinants in normal epithelial cells is the absorption of pathogenic microorganisms such as viruses and bacteria. The metabolic conversion of sialyl 6-sulfo Le^x into cyclic 6-sulfo sLe^x by a calcium dependent enzyme, sialic acid cyclase, occurs as well in nonmalignant colonic epithelia (147), probably driven by epigenetic changes.

Lewis y (Le^y/ CD174, *Fucα1-2Galβ1-4(Fucα1-3)GlcNAc*), another member of the type 1 Lewis glycan structures, has been found in poorly differentiated carcinomas of the liver, colorectum and uterus (148-150). When conjugated to an adjuvant, Le^y has been tested as a potential vaccine in ovarian cancer and has produced good immune response (151).

Involvement of type 2 (*α1-4* fucosylated glycan) Lewis structures, particularly sialyl Lewis a antigen (sLe^a/ *NeuAcα2-3Galβ1-3(Fucα1-4)GlcNAc*), also called clinically carbohydrate-antigen 19-9 (CA19-9), is associated with a variety of gastrointestinal epithelial malignancies and is used as a tumor marker. A gradual increase in the amount of sLe^a was found in colon and rectal tissue during neoplastic transformation and progression (152). Cells derived from human lung cancer were tumorigenic in nude mice only when sLe^a was present on the cell surface (153). Sialyl Le^a is expressed strongly in cancers of the digestive organs and is known to serve as a ligand for vascular E-selectin in haematogenous cancer metastases (125, 154). Clinicopathological studies have revealed the prognostic potential of sLe^a expression in colorectal carcinoma and their association with a high incidence of relapse and reduced survival time (155).

Other non-sialylated versions of Lewis members were observed to be of importance due to their cancer association. Immunohistochemical analysis revealed a significant higher Lewis $^{y/b}$ (Le^y, Le^b) expression level in a large set of breast carcinomas and were associated with their histological grade, tumor type and Nottingham Prognostic Index (156).

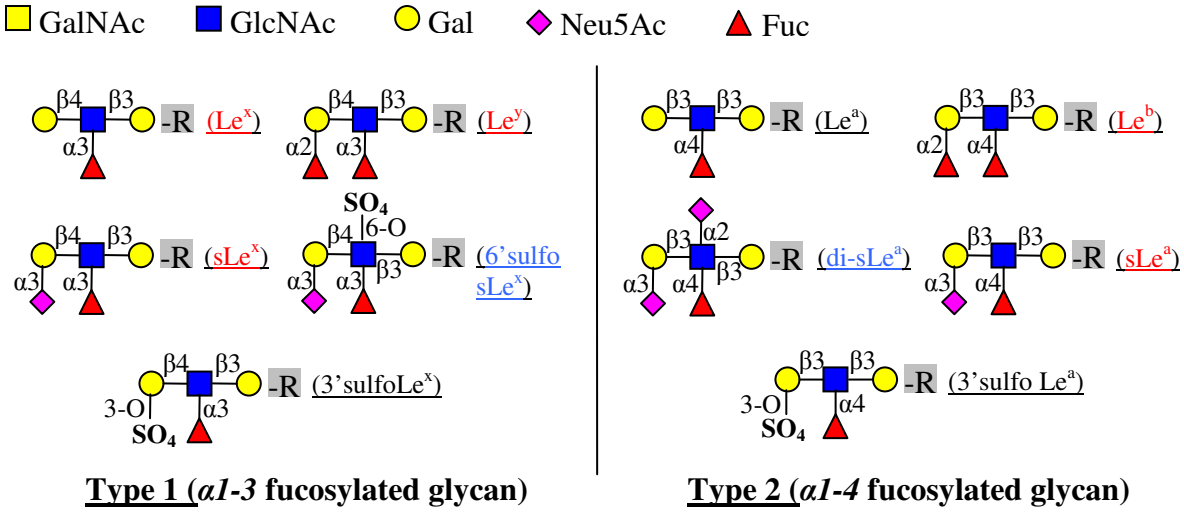


Figure 5 Lewis type 1 and 2 glycan structures.
Lewis glycan structures differ in the linkage of the outermost *Gal* (*β1-3, β1-4*) and in the linkage of the fucose to the *GlcNAc* (*α1-3, α1-4*). Background grayed R, *N*-, *O*- or glycolipid- linkage to biomolecules. Lewis glycan structures associated to cancer (red), disialylated Le^a (di-sLe^a) and 6'sulfo sLe^x (blue) are preferentially expressed in non-malignant epithelial cells.

1.2.2.3 Other Tumor-Associated Carbohydrates

Glycan epitopes such as sTn, Tn and TF in *O*-linked structures, sLe^x, sLe^a and Le^y are well described in different cancer models. This section describes other known tumor-associated carbohydrate structures.

Increased sialylation is one of the most commonly reported alterations in the glycosylation of cancer cell surface glycoproteins and often manifests with an increase in sialylated Lewis or sialylated 'core' *O*-linked structures. Accumulation of *N-glyconeuraminic acid* (*Neu5Gc*) has been reported (1) in melanoma (157) and colorectal cancer (158); on (2) *O*-linked glycans in MUC1 in the breast cancer cell line MCF-7 (159); and on (3) *N*-linked glycans in the metastatic lymphoma cell line MDAY-D2 (160). *Neu5Gc* is produced by another type of sialylation, and differs from *Neu5Ac* by the addition of a single oxygen atom.

A well known phenomenon of glycan association is altered ***N*-glycan branching**. In cancer cells an increase of *β1-6* branching is caused by upregulated gene expression of *GlcNAc* transferase V. This change occurs at the transcriptional level and is induced by a variety of mechanisms including viral and chemical carcinogenesis. High expression of *GlcNAc* transferase V results in an increased metastatic potential of cancer cell lines (109). Increased *β1-6* branching can be measured by a specific lectin known as leukocytic phytohemagglutinin (L-PHA). It does not bind to non-diseased breast epithelial cells, but during the progression to invasive carcinoma, cells display an enhanced ability to L-PHA. Immunohistochemical

staining of L-PHA identified *β1-6* branching as an independent prognostic indicator for poor outcome in primary node-negative tumors (161).

Another family of TACA is comprised of **ganglioside** members. These are sialic acid-bearing glycolipids involved in modulating transmembrane signalling which is essential for tumor cell growth, invasion, and metastasis. The mayor gangliosides GM2 (*Neu5Aca2-3(GalNAcβ1-4)Galβ1-4Glcβ*) and GD3 (*Neu5Aca2-8Neu5Aca2-3Galβ1-4Glcβ*) are known to be expressed by most human cancers of neuro-ectodermal and epithelial origin (162). In melanoma cells, ganglioside expression correlates with metastatic potential. They are also selectively exposed on tumor cells with invasive potential (163, 164). Two randomized trials have used a vaccine against GM2 in advanced stage melanomas (165), bladder cancer (166) and colorectal cancer (167). Patients immunized with GM2 developed naturally occurring IgM antibodies and showed a significant improvement in relapse-free survival (165). A second study confirmed these findings despite use of a different GM2 vaccine (168). GD3 is a minor glycan expressed in most normal tissues (169), its expression has also been described on the surface of normal human peripheral blood T cells (170). GD3 is highly expressed only during development or in pathological conditions like arteriosclerosis and tumors of neuro-ectodermal origin (171).

Globo H (*Fucα1-2Galβ1-3GalNAcβ1-3Galα1-4Galβ1-4Glcβ*), a linear glycosphingolipid, is another well described TACA. A variety of epithelial carcinomas such as colon, ovarian, gastric and lung cancers showed Globo H overexpression (172, 173). High expression was also detected by immunohistochemistry in small cell lung cancers with short overall survival (174). More than 60% of ductal, lobular and tubular breast cancers showed high Gobo H expression whilst no expression could be found in non-epithelial breast tumors (175). Recently, Globo H revealed immune response shown by significantly higher levels of human anti-glycan antibodies (AGA) in breast cancer serum samples than compared to healthy controls (176).

1.2.3 Naturally Occurring Anti-Glycan Antibodies

Apart from secreted glycoproteins most cellular glycoconjugates are found in the plasma membrane, exposing the glycan moiety to the extracellular space. Due to this location and their structural variability, glycans were thought to have a role in cellular “social functions” like recognition (177). It is therefore reasonable to hypothesize that the body has evolved a recognition system to detect and monitor self and pathogenic glycan profiles. In mammalian cells, fungi and parasites carbohydrates or their related glycoconjugates possess

unique structural elements specific to their environment, with some serving as a basis for the recognition of host proteins.

Natural antibody producing B cells, natural killer cells, macrophages and dendritic cells remove all foreign and potentially harmful substances, *e.g.* bacteria and viruses, cellular waste, modified molecules, and most importantly, cancer cells (178). These natural antibodies are made by special types of lymphocytes which play a large role in the humoral immune response. Mature B cells, preferentially B1 or $CD5^+$ cells (179, 180) located in body cavities, are the major source of naturally occurring antibodies. Natural antibodies encoded by their genes during germline maturation (181) are defined as antibodies that circulate in normal individuals in the absence of endogenous or mostly exogenous antigen stimulation. Many natural antibodies are directed to carbohydrate antigens found in normal human tissues and can be called naturally occurring anti-glycan (auto-) antibodies. They are polyreactive (182) belong to immunoglobulin subtypes IgA and IgM, and are present in serum of all individuals even in the absence of previous immunization (183). Thus, they are a part of the innate immune system (184) and probably contribute to the homeostasis of the immune system. Naturally occurring AGA appear in a number of human pathologies, *e.g.* ganglioside GT3 (*Neu5Acα2-8Neu5Acα2-8Neu5Acα2-3Galβ1-4Glc*) in sera of patients with type 1 *diabetes mellitus* (185) or cranial dysfunction (186), in sera of patients with multiple sclerosis containing anti-*Glcα1-4Glcα* (glycogen fragment) antibodies (187), antibodies directed against oligomannose sequences are markers in Crohn's disease (188), and anti-*GlcNAcβ* antibodies in rheumatoid arthritis (189).

Involvement of natural IgM antibodies has been described in the early recognition of exogenous organisms (190), but these antibodies also seem to participate in recognition and elimination of precancerous and cancerous lesions (191, 192). Nearly all investigated tumor-reactive antibodies present in cancer and healthy serum samples are bound to tumor-specific PTM cell surface antigens, later identified as carbohydrate epitopes (192-194). Natural IgM antibodies are also able to induce apoptosis in malignant transformed cells (192, 195). The human mAb SAM-6 was isolated from a gastric cancer patient and exclusively binds to malignant tissues *via* an *O*-linked glycan epitope on receptor GRP78, inducing apoptosis *via* lipid accumulation (196, 197). Another mAb PAM-1, which is directed to an integral membrane receptor homologue of CFR-1, has been described (191, 198, 199) and binds to a tumor-specific *N*-linked glycan located on CFR-1 that is specifically expressed on malignant cells. PAM-1 induces *in vitro* as well as *in vivo* apoptosis probably *via* blockage of CFR-1 (199).

Investigation of intravenous immunoglobulin preparations bound to a printed glycan array (Consortium for Functional Glycomics) containing 377 glycan structures showed that sera pools contain a much wider range of IgG antibodies than expected. Nearly half of the glycans were bound by IgG, with the highest level of binding to structures of human pathogenic and non-pathogenic bacteria. Little to no binding activity to human endogenous glycans, including TACA, was detected (200). Serum levels of natural AGA of IgG subtype bound to TF antigen and were significantly elevated in gastric cancers (201).

In a recent publication extracellular mucin exposed glycans identified MUC1 with aberrant glycosylation patterns in many cancer types. The microarray hydrogel-based platform was developed in order to detect MUC1 *O*-glycopeptide-specific antibodies. Cancer-associated IgG autoantibodies that bound to several *O*-glycopeptide epitopes were found, not including IgG antibody signals to MUC1 peptide epitopes. Human cancer sera (breast, prostate and ovarian) revealed IgG antibody signals directed to Tn-MUC1, sTn-MUC1 and truncated core 3-MUC1, which were not found in healthy control sera (202). Results on core 3 structures are controversial because former studies found higher core 3 expression in normal colon tissue (130, 131).

Globo H is another member of a family of antigenic glycans and is highly expressed on the cell surface of prostate, ovarian, and breast cancer cells (203). Investigations on Globo H and its truncated carbohydrate structures revealed significantly elevated anti-Globo H antibody levels in breast cancer serum samples compared to healthy controls (176). A cancer vaccine based study investigated the potential of Globo H-keyhole limpet hemocyanin (204). They measured a remarkably broad polyclonal antibody activity with a significantly higher IgG level in sera after immunization. Substantial quantities of IgG and IgM antibodies were elicited with a clear indication of a class switch to IgG. Unfortunately, it could not be demonstrated that the IgG antibody levels alone were responsible for the anti-vaccine activity (205).

1.2.4 Distinct Glycosylation in Epithelial Ovarian Cancer (State of Research)

Aberrant glycosylation is a well described hallmark of cancer development and progression. Over 50% of cancers are known to express TACA which includes glycolipids, Lewis blood group antigens, and ‘core’ glycan structures (206). Some indications exist that EOC is also reflected by aberrant glycosylation or distinct glycosylation patterns. Therefore this section highlights the most important findings in EOC and their associated glycan structures.

A recent mass spectrometry study screened oligosaccharides secreted by cancer cell lines and found specific changes in glycosylation without any protein information content. Several oligosaccharides as found in the supernatants of EOC cell lines were concordant with the oligosaccharides detected in serum samples. Moreover, it was found that anionic oligomers were composed of several series of hexuronic acid (108).

In another pilot study, the total serum glycome of patients with advanced EOC was examined. A number of changes were found including increased levels of 'core' fucosylated, agalactosyl biantennary glycans (FA2) and sLe^x. A similar study of glycosylation and glycoprotein levels revealed that acute-phase proteins, haptoglobin, α1-acid glycoprotein, and α1-antitichymotrypsin mostly contain elevated levels of sLe^x. IgG heavy chains from patients contained twice the level of FA2 compared to healthy controls (207). A subsequent study showed that decreased galactosylation and sialylation of IgG heavy chains increased cytotoxity of natural killer cells and complement activation *via* mannose-binding lectin (208).

Even CA125, the best current single biomarker for the detection of EOC contains up to 30% molecular weight in glycan structures, mostly *O*-linked glycans (209). Lewis family carbohydrate structures, including Le^x have been found to be expressed on CA125 (210). A two-step isolation of CA125 from an ovarian cancer cell line and sequencing of oligosaccharides is present in *O*-glycan structures in both core 1 and 2 glycans. An unusual feature was the expression of branched core 1 antennae in the core 2 glycans. It has also been found that CA125 contains *N*-linked glycans, primarily expressing high mannose and complex bisecting type of *N*-linked glycans (211).

Differentiation of EOC histotypes has been studied based on the carbohydrate antigens belonging to O(H) (H type 1, H type 2), Lewis blood group families (Le^a, sLe^a, Le^x, sLe^x, Le^b and Le^y), and the mucin- 'core' family (Tn, sTn and TF) using mAbs on fresh frozen sections. Clear differences in antigen expression were observed between MOC and other histotypes of EOC, particularly as MOC tended to express sTn, Le^a and sLe^a strongly and homogeneously. Strong immunostaining of Le^y and H type 2 antigens were observed in SOC and EnOC only. Based on these findings it has been suggested that MOC and non-MOC tumors are of distinct biology (212).

Tn antigen expression (defined by *Vicia villosa* lectin) was associated with EOC patient survival (213). Its sialylated structure, sTn is elevated in serum samples and correlated with a more aggressive malignant status, metastatic progress and low patient survival (214). Both Tn and sTn expression were also found to be higher in EOC patients than in benign tumors and correlated with shorter overall survival (215). Other new potential TACA were recently

suggested in a study on EOC which revealed that enzyme *GlcNAc* 6-*O*-sulfotransferase-2 is specifically expressed in MOC, ClCOC and SOC. Other new ovarian TACA reported in this study were *6-Su-LacNAc*, *6-Su-Lewis*^{x} and *6-Su-SialylLewis*^{x} (216).

EOC correlated TACA have consistently proven to be promising biomarkers for cancer. Also, the immune response in cancers is clearly and predominantly related to the carbohydrate moiety, independent of their expression as glycoproteins or glycolipids (206). As shown within the results section, using the covalent printed glycan array, the immune response against TACA is indeed far more abundant than previously demonstrated and can be detected in microliters of serum in the form of AGA (176).

1.3 **References**

1. NICER S (2010) (National Institute for Cancer Epidemiology and Registration).
2. Garcia M, Jemal A, & Ward E (2007) Global cancer facts and figures 2007. (American Cancer Society, Atlanta).
3. Wu AH, Pearce CL, Tseng CC, Templeman C, & Pike MC (2009) Markers of inflammation and risk of ovarian cancer in Los Angeles County. *Int J Cancer* 124(6):1409-1415.
4. Hannibal CG, Rossing MA, Wicklund KG, & Cushing-Haugen KL (2008) Analgesic drug use and risk of epithelial ovarian cancer. *Am J Epidemiol* 167(12):1430-1437.
5. Merritt MA, Green AC, Nagle CM, & Webb PM (2008) Talcum powder, chronic pelvic inflammation and NSAIDs in relation to risk of epithelial ovarian cancer. *Int J Cancer* 122(1):170-176.
6. Ness RB*, et al.* (2000) Factors related to inflammation of the ovarian epithelium and risk of ovarian cancer. *Epidemiology* 11(2):111-117.
7. Fathalla MF (1971) Incessant ovulation--a factor in ovarian neoplasia? *Lancet* 2(7716):163.
8. Cramer DW & Welch WR (1983) Determinants of ovarian cancer risk. II. Inferences regarding pathogenesis. *J Natl Cancer Inst* 71(4):717-721.
9. Casagrande JT*, et al.* (1979) "Incessant ovulation" and ovarian cancer. *Lancet* 2(8135):170-173.
10. Risch HA, Weiss NS, Lyon JL, Daling JR, & Liff JM (1983) Events of reproductive life and the incidence of epithelial ovarian cancer. *Am J Epidemiol* 117(2):128-139.
11. Tung KH*, et al.* (2005) Effect of anovulation factors on pre- and postmenopausal ovarian cancer risk: revisiting the incessant ovulation hypothesis. *Am J Epidemiol* 161(4):321-329.
12. Easton DF, Ford D, & Bishop DT (1995) Breast and ovarian cancer incidence in BRCA1-mutation carriers. Breast Cancer Linkage Consortium. *Am J Hum Genet* 56(1):265-271.
13. Antoniou A*, et al.* (2003) Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case Series unselected for family history: a combined analysis of 22 studies. *Am J Hum Genet* 72(5):1117-1130.
14. Shaw PA*, et al.* (2002) Histopathologic features of genetically determined ovarian cancer. *Int J Gynecol Pathol* 21(4):407-411.
15. Lawrenson K & Gayther SA (2009) Ovarian cancer: a clinical challenge that needs some basic answers. *PLoS Med* 6(2):e25.

16. Crawford SC, *et al.* (2005) Does aggressive surgery only benefit patients with less advanced ovarian cancer? Results from an international comparison within the SCOTROC-1 Trial. *J Clin Oncol* 23(34):8802-8811.
17. Hacker NF, Berek JS, Lagasse LD, Nieberg RK, & Elashoff RM (1983) Primary cytoreductive surgery for epithelial ovarian cancer. *Obstet Gynecol* 61(4):413-420.
18. Hennessy BT, Coleman RL, & Markman M (2009) Ovarian cancer. *Lancet* 374(9698):1371-1382.
19. Bhoola S & Hoskins WJ (2006) Diagnosis and management of epithelial ovarian cancer. *Obstet Gynecol* 107(6):1399-1410.
20. Friedlander M, *et al.* (2007) Phase II study of carboplatin followed by sequential gemcitabine and paclitaxel as first-line treatment for advanced ovarian cancer. *Int J Gynecol Cancer* 17(2):350-358.
21. Agarwal R, Linch M, & Kaye SB (2006) Novel therapeutic agents in ovarian cancer. *Eur J Surg Oncol* 32(8):875-886.
22. Coleman BG (1992) Transvaginal sonography of adnexal masses. *Radiol Clin North Am* 30(4):677-691.
23. Granberg S, Wikland M, & Jansson I (1989) Macroscopic characterization of ovarian tumors and the relation to the histological diagnosis: criteria to be used for ultrasound evaluation. *Gynecol Oncol* 35(2):139-144.
24. Bourne T, Campbell S, Steer C, Whitehead MI, & Collins WP (1989) Transvaginal colour flow imaging: a possible new screening technique for ovarian cancer. *Bmj* 299(6712):1367-1370.
25. Lozeau AM & Potter B (2005) Diagnosis and management of ectopic pregnancy. *Am Fam Physician* 72(9):1707-1714.
26. Givens V, Mitchell GE, Harraway-Smith C, Reddy A, & Maness DL (2009) Diagnosis and management of adnexal masses. *Am Fam Physician* 80(8):815-820.
27. Bast RC, Jr., *et al.* (2005) New tumor markers: CA125 and beyond. *Int J Gynecol Cancer* 15 Suppl 3:274-281.
28. Seidman J, Ronnett B, & Kurman R (2002) Pathology of borderline (low malignant potential) ovarian tumours. *Best Practice&Research Clinical Obstetrics and Gynaecology* 16 (4):499-512.
29. Jemal A, *et al.* (2007) Cancer statistics, 2007. *CA Cancer J Clin* 57(1):43-66.
30. Seidman JD, *et al.* (2004) The histologic type and stage distribution of ovarian carcinomas of surface epithelial origin. *Int J Gynecol Pathol* 23(1):41-44.
31. Robboy SJ, Anderson MC, & Russell P (2002) *Pathology of the female reproductive tract* (Churchill Livingstone, London) pp xxi, 929.
32. Levanon K, Crum C, & Drapkin R (2008) New insights into the pathogenesis of serous ovarian cancer and its clinical impact. *J Clin Oncol* 26(32):5284-5293.
33. Malpica A, *et al.* (2004) Grading ovarian serous carcinoma using a two-tier system. *Am J Surg Pathol* 28(4):496-504.
34. Malpica A, *et al.* (2007) Interobserver and intraobserver variability of a two-tier system for grading ovarian serous carcinoma. *Am J Surg Pathol* 31(8):1168-1174.
35. Obata K, *et al.* (1998) Frequent PTEN/MMAC mutations in endometrioid but not serous or mucinous epthelial ovarian tumors. *Cancer Res* 58:2095-2097.
36. Saito M, *et al.* (2000) Allelic imbalance and mutations of the PTEN gene in ovarian cancer. *Int J Cancer* 85(2):160-165.
37. Sato N, *et al.* (2000) Loss of heterozygosity on 10q23.3 and mutation of the tumor suppressor gene PTEN in benign endometrial cyst of the ovary: possible sequence progression from benign endometrial cyst to endometrioid carcinoma and clear cell carcinoma of the ovary. *Cancer Res* 60(24):7052-7056.

38. Thomas EJ & Campbell IG (2000) Molecular genetic defects in endometriosis. *Gynecol Obstet Invest* 50 Suppl 1:44-50.
39. Gilks CB (2010) Molecular abnormalities in ovarian cancer subtypes other than high-grade serous carcinoma. *J Oncol*:740968.
40. Kurman RJ & Shih Ie M (2008) Pathogenesis of ovarian cancer: lessons from morphology and molecular biology and their clinical implications. *Int J Gynecol Pathol* 27(2):151-160.
41. Enomoto T, Weghorst CM, Inoue M, Tanizawa O, & Rice JM (1991) K-ras activation occurs frequently in mucinous adenocarcinomas and rarely in other common epithelial tumors of the human ovary. *Am J Pathol* 139(4):777-785.
42. Gemignani ML*, et al.* (2003) Role of KRAS and BRAF gene mutations in mucinous ovarian carcinoma. *Gynecol Oncol* 90:378-381.
43. Mayr D, Hirschmann A, Lohrs U, & Diebold J (2006) KRAS and BRAF mutations in ovarian tumors: a comprehensive study of invasive carcinomas, borderline tumors and extraovarian implants. *Gynecol Oncol* 103(3):883-887.
44. Santini D*, et al.* (1989) Brenner tumor of the ovary: a correlative histologic, histochemical, immunohistochemical, and ultrastructural investigation. *Hum Pathol* 20(8):787-795.
45. Crum CP*, et al.* (2007) Lessons from BRCA: the tubal fimbria emerges as an origin for pelvic serous cancer. *Clin Med Res* 5(1):35-44.
46. Auersperg N, Wong AS, Choi KC, Kang SK, & Leung PC (2001) Ovarian surface epithelium: biology, endocrinology, and pathology. *Endocr Rev* 22(2):255-288.
47. Guioli S, Sekido R, & Lovell-Badge R (2007) The origin of the Mullerian duct in chick and mouse. *Dev Biol* 302(2):389-398.
48. Kobayashi A & Behringer RR (2003) Developmental genetics of the female reproductive tract in mammals. *Nat Rev Genet* 4(12):969-980.
49. Feeley KM & Wells M (2001) Precursor lesions of ovarian epithelial malignancy. *Histopathology* 38(2):87-95.
50. Naora H (2007) The heterogeneity of epithelial ovarian cancers: reconciling old and new paradigms. *Expert Rev Mol Med* 9(13):1-12.
51. Cheng W, Liu J, Yoshida H, Rosen D, & Naora H (2005) Lineage infidelity of epithelial ovarian cancers is controlled by HOX genes that specify regional identity in the reproductive tract. *Nat Med* 11(5):531-537.
52. Hennessy BT & Mills GB (2006) Ovarian cancer: homeobox genes, autocrine/paracrine growth, and kinase signaling. *Int J Biochem Cell Biol* 38(9):1450-1456.
53. Dubeau L (1999) The cell of origin of ovarian epithelial tumors and the ovarian surface epithelium dogma: does the emperor have no clothes? *Gynecol Oncol* 72(3):437-442.
54. Liu J*, et al.* (2004) A genetically defined model for human ovarian cancer. *Cancer Res* 64(5):1655-1663.
55. Connolly DC*, et al.* (2003) Female mice chimeric for expression of the simian virus 40 TAg under control of the MISIIR promoter develop epithelial ovarian cancer. *Cancer Res* 63(6):1389-1397.
56. Flesken-Nikitin A, Choi KC, Eng JP, Shmidt EN, & Nikitin AY (2003) Induction of carcinogenesis by concurrent inactivation of p53 and Rb1 in the mouse ovarian surface epithelium. *Cancer Res* 63(13):3459-3463.
57. Kim HS & Song YS (2009) International Federation of Gynecology and Obstetrics (FIGO) staging system revised: what should be considered critically for gynecologic cancer? *J Gynecol Oncol* 20(3):135-136.

58. Jacob F, Goldstein DR, Fink D, & Heinzelmann-Schwarz V (2009) Proteogenomic studies in epithelial ovarian cancer: established knowledge and future needs. *Biomarkers Med.* 3(6):743-756.
59. Singer G, Shih IM, Truskinovsky A, Umudum H, & Kurman RJ (2003) Mutational analysis of K-ras segregates ovarian serous carcinomas into two types: invasive MPSC (low-grade tumor) and conventional serous carcinoma (high-grade tumor). *Int J Gynecol Pathol* 22:37-41.
60. Singer G*, et al.* (2005) Patterns of p53 mutations separate ovarian serous borderline tumors and low- and high-grade carcinomas and provide support for a new model of ovarian carcinogenesis: a mutational analysis with immunohistochemical correlation. *Am J Surg Pathol* 29(2):218-224.
61. Shih Ie M & Kurman RJ (2004) Ovarian tumorigenesis: a proposed model based on morphological and molecular genetic analysis. *Am J Pathol* 164:1511-1518.
62. Crum CP*, et al.* (2007) The distal fallopian tube: a new model for pelvic serous carcinogenesis. *Curr Opin Obstet Gynecol* 19(1):3-9.
63. Jarboe E*, et al.* (2008) Serous carcinogenesis in the fallopian tube: a descriptive classification. *Int J Gynecol Pathol* 27(1):1-9.
64. Kindelberger DW*, et al.* (2007) Intraepithelial carcinoma of the fimbria and pelvic serous carcinoma: Evidence for a causal relationship. *Am J Surg Pathol* 31(2):161-169.
65. Medeiros F*, et al.* (2006) The tubal fimbria is a preferred site for early adenocarcinoma in women with familial ovarian cancer syndrome. *Am J Surg Pathol* 30(2):230-236.
66. Shih Ie M & Kurman RJ (2005) Molecular pathogenesis of ovarian borderline tumors: new insights and old challenges. *Clin Cancer Res* 11(20):7273-7279.
67. Cho KR (2009) Ovarian cancer update: lessons from morphology, molecules, and mice. *Arch Pathol Lab Med* 133(11):1775-1781.
68. Marchini S*, et al.* (2008) Analysis of gene expression in early-stage ovarian cancer. *Clin Cancer Res* 14(23):7850-7860.
69. Purdie DM, Webb PM, Siskind V, Bain CJ, & Green AC (2003) The different etiologies of mucinous and nonmucinous epithelial ovarian cancers. *Gynecol Oncol* 88(1 Pt 2):S145-148.
70. Jordan SJ, Whiteman DC, Purdie DM, Green AC, & Webb PM (2006) Does smoking increase risk of ovarian cancer? A systematic review. *Gynecol Oncol* 103(3):1122-1129.
71. Pal T*, et al.* (2005) BRCA1 and BRCA2 mutations account for a large proportion of ovarian carcinoma cases. *Cancer* 104(12):2807-2816.
72. Heinzelmann-Schwarz VA*, et al.* (2006) A distinct molecular profile associated with mucinous epithelial ovarian cancer. *Br J Cancer* 94(6):904-913.
73. Cai KQ*, et al.* (2009) Loss of GATA4 and GATA6 expression specifies ovarian cancer histological subtypes and precedes neoplastic transformation of ovarian surface epithelia. *PLoS One* 4(7):e6454.
74. Bast RC, Jr.*, et al.* (1981) Reactivity of a monoclonal antibody with human ovarian carcinoma. *J Clin Invest* 68(5):1331-1337.
75. Bast RC, Jr.*, et al.* (1983) A radioimmunoassay using a monoclonal antibody to monitor the course of epithelial ovarian cancer. *N Engl J Med* 309(15):883-887.
76. Bast RC, Jr.*, et al.* (1998) CA 125: the past and the future. *Int J Biol Markers* 13(4):179-187.
77. Duffy MJ*, et al.* (2005) CA125 in ovarian cancer: European Group on Tumor Markers guidelines for clinical use. *Int J Gynecol Cancer* 15(5):679-691.
78. Hough C*, et al.* (2000) Large-scale serial analysis of gene expression reveals genes differentially expressed in ovarian cancer. *Cancer Research* 60:6281-6287.

79. Schummer M, *et al.* (1999) Comparative hybridization of an array of 21,500 ovarian cDNAs for the discovery of genes overexpressed in ovarian carcinomas. *Gene* 238(2):375-385.
80. Mok SC, *et al.* (2001) Prostasin, a potential serum marker for ovarian cancer: identification through microarray technology. *J Natl Cancer Inst* 93(19):1458-1464.
81. Kim JH, *et al.* (2002) Osteopontin as a potential diagnostic biomarker for ovarian cancer. *Jama* 287(13):1671-1679.
82. Simon I, *et al.* (2006) B7-h4 is a novel membrane-bound protein and a candidate serum and tissue biomarker for ovarian cancer. *Cancer Res* 66(3):1570-1575.
83. Simon I, *et al.* (2007) B7-H4 is over-expressed in early-stage ovarian cancer and is independent of CA125 expression. *Gynecol Oncol* 106(2):334-341.
84. Chang K & Pastan I (1996) Molecular cloning of mesothelin, a differentiation antigen present on mesothelium, mesotheliomas, and ovarian cancers. *Proc Natl Acad Sci U S A* 93(1):136-140.
85. Hellstrom I & Hellstrom KE (2008) SMRP and HE4 as biomarkers for ovarian carcinoma when used alone and in combination with CA125 and/or each other. *Adv Exp Med Biol* 622:15-21.
86. Cheng WF, *et al.* (2009) High mesothelin correlates with chemoresistance and poor survival in epithelial ovarian carcinoma. *Br J Cancer* 100(7):1144-1153.
87. Huang CY, *et al.* (2006) Serum mesothelin in epithelial ovarian carcinoma: a new screening marker and prognostic factor. *Anticancer Res* 26(6C):4721-4728.
88. Baron AT, *et al.* (1999) Serum sErbB1 and epidermal growth factor levels as tumor biomarkers in women with stage III or IV epithelial ovarian cancer. *Cancer Epidemiol Biomarkers Prev* 8(2):129-137.
89. Baron AT, *et al.* (2003) Soluble epidermal growth factor receptor (sEGFR/sErbB1) as a potential risk, screening, and diagnostic serum biomarker of epithelial ovarian cancer. *Cancer Epidemiol Biomarkers Prev* 12(2):103-113.
90. Moore RG, *et al.* (2008) The use of multiple novel tumor biomarkers for the detection of ovarian carcinoma in patients with a pelvic mass. *Gynecol Oncol* 108(2):402-408.
91. Moore RG, *et al.* (2009) A novel multiple marker bioassay utilizing HE4 and CA125 for the prediction of ovarian cancer in patients with a pelvic mass. *Gynecol Oncol* 112(1):40-46.
92. Mor G, *et al.* (2005) Serum protein markers for early detection of ovarian cancer. *Proc Natl Acad Sci U S A* 102(21):7677-7682.
93. Tsigkou A, *et al.* (2007) Total inhibin is a potential serum marker for epithelial ovarian cancer. *J Clin Endocrinol Metab* 92(7):2526-2531.
94. Kishi T, *et al.* (2003) Human kallikrein 8, a novel biomarker for ovarian carcinoma. *Cancer Res* 63(11):2771-2774.
95. Scholler N, *et al.* (1999) Soluble member(s) of the mesothelin/megakaryocyte potentiating factor family are detectable in sera from patients with ovarian carcinoma. *Proc Natl Acad Sci U S A* 96(20):11531-11536.
96. Kuk C, *et al.* (Nidogen-2: a new serum biomarker for ovarian cancer. *Clin Biochem* 43(4-5):355-361.
97. Visintin I, *et al.* (2008) Diagnostic markers for early detection of ovarian cancer. *Clin Cancer Res* 14(4):1065-1072.
98. Bertenshaw GP, *et al.* (2008) Multianalyte profiling of serum antigens and autoimmune and infectious disease molecules to identify biomarkers dysregulated in epithelial ovarian cancer. *Cancer Epidemiol Biomarkers Prev* 17(10):2872-2881.
99. Petricoin EF, *et al.* (2002) Use of proteomic patterns in serum to identify ovarian cancer. *Lancet* 359(9306):572-577.

100. Rai AJ, *et al.* (2002) Proteomic approaches to tumor marker discovery. *Arch Pathol Lab Med* 126(12):1518-1526.
101. Moshkovskii SA, *et al.* (2005) Ovarian cancer marker of 11.7 kDa detected by proteomics is a serum amyloid A1. *Proteomics* 5(14):3790-3797.
102. Ogata Y, *et al.* (2006) Elevated levels of phosphorylated fibrinogen-alpha-isoforms and differential expression of other post-translationally modified proteins in the plasma of ovarian cancer patients. *J Proteome Res* 5(12):3318-3325.
103. Lopez MF, *et al.* (2007) A novel, high-throughput workflow for discovery and identification of serum carrier protein-bound peptide biomarker candidates in ovarian cancer samples. *Clin Chem* 53(6):1067-1074.
104. Lokshin AE, *et al.* (2006) Circulating IL-8 and anti-IL-8 autoantibody in patients with ovarian cancer. *Gynecol Oncol* 102(2):244-251.
105. Gagnon A, *et al.* (2008) Use of a combination of approaches to identify and validate relevant tumor-associated antigens and their corresponding autoantibodies in ovarian cancer patients. *Clin Cancer Res* 14(3):764-771.
106. Lu H, Goodell V, & Disis ML (2008) Humoral immunity directed against tumor-associated antigens as potential biomarkers for the early diagnosis of cancer. *J Proteome Res* 7(4):1388-1394.
107. Tsai-Turton M, *et al.* (2009) p53 autoantibodies, cytokine levels and ovarian carcinogenesis. *Gynecol Oncol* 114(1):12-17.
108. An HJ, *et al.* (2006) Profiling of glycans in serum for the discovery of potential biomarkers for ovarian cancer. *J Proteome Res* 5(7):1626-1635.
109. Varki A, *et al.* (1999) *Essentials of Glycobiology* (Cold Spring Harbor Labartory Press, Plainview (NY)) p p.41.
110. Rambourg A & Leblond CP (1967) Electron microscope observations on the carbohydrate-rich cell coat present at the surface of cells in the rat. *J Cell Biol* 32(1):27-53.
111. Apweiler R, Hermjakob H, & Sharon N (1999) On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim Biophys Acta* 1473(1):4-8.
112. Brooks SA, *et al.* (2008) Altered glycosylation of proteins in cancer: what is the potential for new anti-tumour strategies. *Anticancer Agents Med Chem* 8(1):2-21.
113. Spiro RG (2002) Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology* 12(4):43R-56R.
114. Turnbull JE & Field RA (2007) Emerging glycomics technologies. *Nat Chem Biol* 3(2):74-77.
115. Hart GW (1992) Glycosylation. *Curr Opin Cell Biol* 4(6):1017-1023.
116. Seitz O (2000) Glycopeptide synthesis and the effects of glycosylation on protein structure and activity. *Chembiochem* 1(4):214-246.
117. Bennett EP, *et al.* (1998) Genomic organization and chromosomal localization of three members of the UDP-N-acetylgalactosamine: polypeptide N-acetylgalactosaminyltransferase family. *Glycobiology* 8(6):547-555.
118. Chandrasekaran A, *et al.* (2008) Glycan topology determines human adaptation of avian H5N1 virus hemagglutinin. *Nat Biotechnol* 26(1):107-113.
119. Guo BP, Norris SJ, Rosenberg LC, & Hook M (1995) Adherence of Borrelia burgdorferi to the proteoglycan decorin. *Infect Immun* 63(9):3467-3472.
120. Hoke DE, Egan S, Cullen PA, & Adler B (2008) LipL32 is an extracellular matrix-interacting protein of Leptospira spp. and Pseudoalteromonas tunicata. *Infect Immun* 76(5):2063-2069.

121. Amano K, Hayashi S, Kubota T, Fujii N, & Yokota S (1997) Reactivities of Lewis antigen monoclonal antibodies with the lipopolysaccharides of Helicobacter pylori strains isolated from patients with gastroduodenal diseases in Japan. *Clin Diagn Lab Immunol* 4(5):540-544.
122. Felner KM, Dinter A, Cartron JP, & Berger EG (1998) Repressed beta-1,3-galactosyltransferase in the Tn syndrome. *Biochim Biophys Acta* 1406(1):115-125.
123. Piller F, Le Deist F, Weinberg KI, Parkman R, & Fukuda M (1991) Altered O-glycan synthesis in lymphocytes from patients with Wiskott-Aldrich syndrome. *J Exp Med* 173(6):1501-1510.
124. Horiuchi K*, et al.* (2005) Three-dimensional ultrastructure of the brush border glycocalyx in the mouse small intestine: a high resolution scanning electron microscopic study. *Arch Histol Cytol* 68(1):51-56.
125. Hakomori S (2002) Glycosylation defining cancer malignancy: new wine in an old bottle. *PNAS* 99(16):10231-10233.
126. Julien S*, et al.* (2001) Expression of sialyl-Tn antigen in breast cancer cells transfected with the human CMP-Neu5Ac: GalNAc alpha2,6-sialyltransferase (ST6GalNac I) cDNA. *Glycoconj J* 18(11-12):883-893.
127. Marcos NT*, et al.* (2004) Role of the human ST6GalNAc-I and ST6GalNAc-II in the synthesis of the cancer-associated sialyl-Tn antigen. *Cancer Res* 64(19):7050-7057.
128. Kinney AY*, et al.* (1997) The prognostic significance of sialyl-Tn antigen in women treated with breast carcinoma treated with adjuvant chemotherapy. *Cancer* 80(12):2240-2249.
129. Baldus SE*, et al.* (2000) Thomsen-Friedenreich antigen presents as a prognostic factor in colorectal carcinoma: A clinicopathologic study of 264 patients. *Cancer* 88(7):1536-1543.
130. Brockhausen I (1999) Pathways of O-glycan biosynthesis in cancer cells. *Biochim Biophys Acta* 1473(1):67-95.
131. Capon C, Maes E, Michalski JC, Leffler H, & Kim YS (2001) Sd(a)-antigen-like structures carried on core 3 are prominent features of glycans from the mucin of normal human descending colon. *Biochem J* 358(Pt 3):657-664.
132. Kurosaka A*, et al.* (1983) Structures of the major oligosaccharides from a human rectal adenocarcinoma glycoprotein. *J Biol Chem* 258(19):11594-11598.
133. Feeney J, Frenkiel TA, & Hounsell EF (1986) Complete 1H-n.m.r. assignments for two core-region oligosaccharides of human meconium glycoproteins, using 1D and 2D methods at 500 MHz. *Carbohydr Res* 152:63-72.
134. Robbe C, Capon C, Coddeville B, & Michalski JC (2004) Structural diversity and specific distribution of O-glycans in normal human mucins along the intestinal tract. *Biochem J* 384(Pt 2):307-316.
135. Ravn V & Dabelsteen E (2000) Tissue distribution of histo-blood group antigens. *Apmis* 108(1):1-28.
136. Boren T, Falk P, Roth KA, Larson G, & Normark S (1993) Attachment of Helicobacter pylori to human gastric epithelium mediated by blood group antigens. *Science* 262(5141):1892-1895.
137. Hilton E, Chandrasekaran V, Rindos P, & Isenberg HD (1995) Association of recurrent candidal vaginitis with inheritance of Lewis blood group antigens. *J Infect Dis* 172(6):1616-1619.
138. Stapleton A, Nudelman E, Clausen H, Hakomori S, & Stamm WE (1992) Binding of uropathogenic Escherichia coli R45 to glycolipids extracted from vaginal epithelial cells is dependent on histo-blood group secretor status. *J Clin Invest* 90(3):965-972.

139. Hein HO, Sorensen H, Suadicani P, & Gyntelberg F (1992) The Lewis blood group--a new genetic marker of ischaemic heart disease. *J Intern Med* 232(6):481-487.
140. Hirschberg CB (2001) Golgi nucleotide sugar transport and leukocyte adhesion deficiency II. *J Clin Invest* 108(1):3-6.
141. Oriol R, Opelz G, Chun C, & Terasaki PI (1980) The Lewis system and kidney transplantation. *Transplantation* 29(5):397-400.
142. Shi ZR, McIntyre LJ, Knowles BB, Solter D, & Kim YS (1984) Expression of a carbohydrate differentiation antigen, stage-specific embryonic antigen 1, in human colonic adenocarcinoma. *Cancer Res* 44(3):1142-1147.
143. Cordon-Cardo C*, et al.* (1989) Blood group-related antigens in human kidney: modulation of Lewis determinants in renal cell carcinoma. *Cancer Res* 49(1):212-218.
144. Shirahama T, Ikoma M, Muramatsu H, Muramatsu T, & Ohi Y (1992) Reactivity to fucose-binding proteins of Lotus tetragonolobus correlates with metastatic phenotype of transitional cell carcinoma of the bladder. *J Urol* 147(6):1659-1664.
145. Tabares G*, et al.* (2006) Different glycan structures in prostate-specific antigen from prostate cancer sera in relation to seminal plasma PSA. *Glycobiology* 16(2):132-145.
146. Wang QY, Wu SL, Chen JH, Liu F, & Chen HL (2003) Expressions of Lewis antigens in human non-small cell pulmonary cancer and primary liver cancer with different pathological conditions. *J Exp Clin Cancer Res* 22(3):431-440.
147. Izawa M*, et al.* (2000) Expression of sialyl 6-sulfo Lewis X is inversely correlated with conventional sialyl Lewis X expression in human colorectal cancer. *Cancer Res* 60(5):1410-1416.
148. Abe K, Hakomori S, & Ohshiba S (1986) Differential expression of difucosyl type 2 chain (LeY) defined by monoclonal antibody AH6 in different locations of colonic epithelia, various histological types of colonic polyps, and adenocarcinomas. *Cancer Res* 46(5):2639-2644.
149. Umezaki K, Taga K, Nakajima T, Sawaragi I, & Okamura A (1995) [Expression of Le(y) antigen in adenocarcinoma and its related lesions of the human uterine cervix]. *Nippon Sanka Fujinka Gakkai Zasshi* 47(9):925-930.
150. Wakabayashi M*, et al.* (1995) Lewis Y antigen expression in hepatocellular carcinoma. An immunohistochemical study. *Cancer* 75(12):2827-2835.
151. Sabbatini PJ*, et al.* (2000) Immunization of ovarian cancer patients with a synthetic Lewis(y)-protein conjugate vaccine: a phase 1 trial. *Int J Cancer* 87(1):79-85.
152. Gong E*, et al.* (1985) Expression of carbohydrate antigen 19-9 and stage-specific embryonic antigen 1 in nontumorous and tumorous epithelia of the human colon and rectum. *J Natl Cancer Inst* 75(3):447-454.
153. Pettijohn DE, Pfenninger O, Brown J, Duke R, & Olsson L (1988) Tumorigenic human squamous lung cancer cells have defined cell surface carbohydrates that are absent from nontumorigenic cells. *Proc Natl Acad Sci U S A* 85(3):802-806.
154. Hakomori S (1996) Tumor malignancy defined by aberrant glycosylation and sphingo(glyco)lipid metabolism. *Cancer Res* 56(23):5309-5318.
155. Nakayama T, Watanabe M, Katsumata T, Teramoto T, & Kitajima M (1995) Expression of sialyl Lewis(a) as a new prognostic factor for patients with advanced colorectal carcinoma. *Cancer* 75(8):2051-2056.
156. Madjd Z*, et al.* (2005) High expression of Lewis y/b antigens is associated with decreased survival in lymph node negative breast carcinomas. *Breast Cancer Res* 7(5):R780-787.
157. Furukawa K, Yamaguchi H, Oettgen HF, Old LJ, & Lloyd KO (1989) Two human monoclonal antibodies reacting with the major gangliosides of human melanomas and

comparison with corresponding mouse monoclonal antibodies. *Cancer Res* 49(1):191-196.
158. Higashi H*, et al.* (1985) Characterization of N-glycolylneuraminic acid-containing gangliosides as tumor-associated Hanganutziu-Deicher antigen in human colon cancer. *Cancer Res* 45(8):3796-3802.
159. Devine PL*, et al.* (1991) The breast tumor-associated epitope defined by monoclonal antibody 3E1.2 is an O-linked mucin carbohydrate containing N-glycolylneuraminic acid. *Cancer Res* 51(21):5826-5836.
160. Takano R, Muchmore E, & Dennis JW (1994) Sialylation and malignant potential in tumour cell glycosylation mutants. *Glycobiology* 4(5):665-674.
161. Handerson T, Camp R, Harigopal M, Rimm D, & Pawelek J (2005) Beta1,6-branched oligosaccharides are increased in lymph node metastases and predict poor outcome in breast carcinoma. *Clin Cancer Res* 11(8):2969-2973.
162. Nakamura K, Tanaka Y, Shitara K, & Hanai N (2001) Construction of humanized anti-ganglioside monoclonal antibodies with potent immune effector functions. *Cancer Immunol Immunother* 50(5):275-284.
163. Distler U*, et al.* (2008) Tumor-associated CD75s- and iso-CD75s-gangliosides are potential targets for adjuvant therapy in pancreatic cancer. *Mol Cancer Ther* 7(8):2464-2475.
164. Fredman P, Hedberg K, & Brezicka T (2003) Gangliosides as therapeutic targets for cancer. *BioDrugs* 17(3):155-167.
165. Livingston PO*, et al.* (1994) Improved survival in stage III melanoma patients with GM2 antibodies: a randomized trial of adjuvant vaccination with GM2 ganglioside. *J Clin Oncol* 12(5):1036-1044.
166. Wishahi MM, Ismail IM, & el-Sherbini M (1994) Immunotherapy with bacille Calmette-Guerin in patients with superficial transitional cell carcinoma of the bladder associated with bilharziasis. *Br J Urol* 73(6):649-654.
167. Mosolits S, Nilsson B, & Mellstedt H (2005) Towards therapeutic vaccines for colorectal carcinoma: a review of clinical trials. *Expert Rev Vaccines* 4(3):329-350.
168. Kirkwood JM*, et al.* (2001) High-dose interferon alfa-2b does not diminish antibody response to GM2 vaccination in patients with resected melanoma: results of the Multicenter Eastern Cooperative Oncology Group Phase II Trial E2696. *J Clin Oncol* 19(5):1430-1436.
169. Dyatlovitskaya EV & Bergelson LD (1987) Glycosphingolipids and antitumor immunity. *Biochim Biophys Acta* 907(2):125-143.
170. Merritt WD, Taylor BJ, Der-Minassian V, & Reaman GH (1996) Coexpression of GD3 ganglioside with CD45RO in resting and activated human T lymphocytes. *Cell Immunol* 173(1):131-148.
171. Malisan F & Testi R (2002) GD3 ganglioside and apoptosis. *Biochim Biophys Acta* 1585(2-3):179-187.
172. Bremer EG*, et al.* (1984) Characterization of a glycosphingolipid antigen defined by the monoclonal antibody MBr1 expressed in normal and neoplastic epithelial cells of human mammary gland. *J Biol Chem* 259(23):14773-14777.
173. Menard S, Tagliabue E, Canevari S, Fossati G, & Colnaghi MI (1983) Generation of monoclonal antibodies reacting with normal and cancer cells of human breast. *Cancer Res* 43(3):1295-1300.
174. Martignone S*, et al.* (1989) Relationship between CaMBr1 expression and tumor progression in small cell lung carcinomas. *Tumori* 75(4):373-377.

175. Mariani-Costantini R, *et al.* (1984) Reactivity of a monoclonal antibody with tissues and tumors from the human breast. Immunohistochemical localization of a new antigen and clinicopathologic correlations. *Am J Pathol* 115(1):47-56.
176. Wang CC, *et al.* (2008) Glycan microarray of Globo H and related structures for quantitative analysis of breast cancer. *Proc Natl Acad Sci U S A* 105(33):11661-11666.
177. Hakomori S & Igarashi Y (1995) Functional role of glycosphingolipids in cell recognition and signaling. *J Biochem* 118(6):1091-1103.
178. Vollmers HP & Brandlein S (2009) Natural antibodies and cancer. *N Biotechnol* 25(5):294-298.
179. Baumgarth N, *et al.* (2000) B-1 and B-2 cell-derived immunoglobulin M antibodies are nonredundant components of the protective response to influenza virus infection. *J Exp Med* 192(2):271-280.
180. Casali P & Schettino EW (1996) Structure and function of natural antibodies. *Curr Top Microbiol Immunol* 210:167-179.
181. Coutinho A, Kazatchkine MD, & Avrameas S (1995) Natural autoantibodies. *Curr Opin Immunol* 7(6):812-818.
182. Avrameas S (1991) Natural autoantibodies: from 'horror autotoxicus' to 'gnothi seauton'. *Immunol Today* 12(5):154-159.
183. Shoenfeld Y & Isenberg DA (1989) The mosaic of autoimmunity. *Immunol Today* 10(4):123-126.
184. Vollmers HP & Brandlein S (2007) Natural antibodies and cancer. *J Autoimmun* 29(4):295-302.
185. Gillard BK, Thomas JW, Nell LJ, & Marcus DM (1989) Antibodies against ganglioside GT3 in the sera of patients with type I diabetes mellitus. *J Immunol* 142(11):3826-3832.
186. Koga M, Yuki N, Ariga T, & Hirata K (1999) Antibodies to GD3, GT3, and O-acetylated species in Guillain-Barre and Fisher's syndromes: their association with cranial nerve dysfunction. *J Neurol Sci* 164(1):50-55.
187. Dotan N, Altstock RT, Schwarz M, & Dukler A (2006) Anti-glycan antibodies as biomarkers for diagnosis and prognosis. *Lupus* 15(7):442-450.
188. Standaert-Vitse A, *et al.* (2006) Candida albicans is an immunogen for anti-Saccharomyces cerevisiae antibody markers of Crohn's disease. *Gastroenterology* 130(6):1764-1775.
189. Parekh RB, *et al.* (1985) Association of rheumatoid arthritis and primary osteoarthritis with changes in the glycosylation pattern of total serum IgG. *Nature* 316(6027):452-457.
190. Ulvestad E, *et al.* (2001) Diagnostic and biological significance of anti-p41 IgM antibodies against Borrelia burgdorferi. *Scand J Immunol* 53(4):416-421.
191. Brandlein S, *et al.* (2004) PAM-1, a natural human IgM antibody as new tool for detection of breast and prostate precursors. *Hum Antibodies* 13(4):97-104.
192. Brandlein S, *et al.* (2003) Natural IgM antibodies and immunosurveillance mechanisms against epithelial cancer cells in humans. *Cancer Res* 63(22):7995-8005.
193. Hensel F, *et al.* (1999) Characterization of glycosylphosphatidylinositol-linked molecule CD55/decay-accelerating factor as the receptor for antibody SC-1-induced apoptosis. *Cancer Res* 59(20):5299-5306.
194. Rauschert N, *et al.* (2008) A new tumor-specific variant of GRP78 as target for antibody-based therapy. *Lab Invest* 88(4):375-386.
195. Vollmers HP & Brandlein S (2005) Death by stress: natural IgM-induced apoptosis. *Methods Find Exp Clin Pharmacol* 27(3):185-191.

196. Pohle T, Brandlein S, Ruoff N, Muller-Hermelink HK, & Vollmers HP (2004) Lipoptosis: tumor-specific cell death by antibody-induced intracellular lipid accumulation. *Cancer Res* 64(11):3900-3906.
197. Brandlein S*, et al.* (2007) The human IgM antibody SAM-6 induces tumor-specific apoptosis with oxidized low-density lipoprotein. *Mol Cancer Ther* 6(1):326-333.
198. Hensel F*, et al.* (2001) A novel proliferation-associated variant of CFR-1 defined by a human monoclonal antibody. *Lab Invest* 81(8):1097-1108.
199. Brandlein S*, et al.* (2003) Cysteine-rich fibroblast growth factor receptor 1, a new marker for precancerous epithelial lesions defined by the human monoclonal antibody PAM-1. *Cancer Res* 63(9):2052-2061.
200. von Gunten S*, et al.* (2009) Intravenous immunoglobulin contains a broad repertoire of anticarbohydrate antibodies that is not restricted to the IgG2 subclass. *J Allergy Clin Immunol* 123(6):1268-1276 e1215.
201. Kurtenkov O*, et al.* (2007) Humoral immune response to MUC1 and to the Thomsen-Friedenreich (TF) glycotope in patients with gastric cancer: relation to survival. *Acta Oncol* 46(3):316-323.
202. Wandall HH*, et al.* (2010) Cancer biomarkers defined by autoantibody signatures to aberrant O-glycopeptide epitopes. *Cancer Res* 70(4):1306-1313.
203. Huang CY*, et al.* (2006) Carbohydrate microarray for profiling the antibodies interacting with Globo H tumor antigen. *Proc Natl Acad Sci U S A* 103(1):15-20.
204. Wang ZG*, et al.* (2000) Polyclonal antibodies from patients immunized with a globo H-keyhole limpet hemocyanin vaccine: isolation, quantification, and characterization of immune responses by using totally synthetic immobilized tumor antigens. *Proc Natl Acad Sci U S A* 97(6):2719-2724.
205. Gilboa E (2004) The promise of cancer vaccines. *Nat Rev Cancer* 4(5):401-411.
206. Slovin SF, Keding SJ, & Ragupathi G (2005) Carbohydrate vaccines as immunotherapy for cancer. *Immunol Cell Biol* 83(4):418-428.
207. Saldova R*, et al.* (2007) Ovarian cancer is associated with changes in glycosylation in both acute-phase proteins and IgG. *Glycobiology* 17(12):1344-1356.
208. Saldova R, Wormald MR, Dwek RA, & Rudd PM (2008) Glycosylation changes on serum glycoproteins in ovarian cancer may contribute to disease pathogenesis. *Dis Markers* 25(4-5):219-232.
209. O'Brien TJ*, et al.* (2001) The CA 125 gene: an extracellular superstructure dominated by repeat sequences. *Tumour Biol* 22(6):348-366.
210. Lloyd KO, Yin BW, & Kudryashov V (1997) Isolation and characterization of ovarian cancer antigen CA 125 using a new monoclonal antibody (VK-8): identification as a mucin-type molecule. *Int J Cancer* 71(5):842-850.
211. Kui Wong N*, et al.* (2003) Characterization of the oligosaccharides associated with the human ovarian tumor marker CA125. *J Biol Chem* 278(31):28619-28634.
212. Federici MF, Kudryashov V, Saigo PE, Finstad CL, & Lloyd KO (1999) Selection of carbohydrate antigens in human epithelial ovarian cancers as targets for immunotherapy: serous and mucinous tumors exhibit distinctive patterns of expression. *Int J Cancer* 81(2):193-198.
213. Hakomori S (2001) Tumor-associated carbohydrate antigens defining tumor malignancy: basis for development of anti-cancer vaccines. *Adv Exp Med Biol* 491:369-402.
214. Kobayashi H, Terao T, & Kawashima Y (1992) Sialyl Tn as a prognostic marker in epithelial ovarian cancer. *Br J Cancer* 66(5):984-985.

215. Ghazizadeh M, Ogawa H, Sasaki Y, Araki T, & Aihara K (1997) Mucin carbohydrate antigens (T, Tn, and sialyl-Tn) in human ovarian carcinomas: relationship with histopathology and prognosis. *Hum Pathol* 28(8):960-966.
216. Kanoh A*, et al.* (2006) Ectopic expression of N-acetylglucosamine 6-O-sulfotransferase 2 in chemotherapy-resistant ovarian adenocarcinomas. *Glycoconj J* 23(5-6):453-460.

2 Thesis Aim and Structure

2.1 Background

Due to a poor prognosis in ovarian cancer patients our translational research group focuses on the detection of early cancer biomarkers in this devastating disease. In this PhD thesis we concentrate on past, present and future high-throughput technologies for the detection of ovarian cancer. After a primary literature analysis we have performed a world-wide new discovery approach facilitating a printed glycan array in non-mucinous borderline tumors and ovarian cancer patients compared to healthy controls. Our aim was to investigate the immune response in the form of anti-glycan antibodies towards ovarian cancer-associated glycans in order to improve the detection rate of this disease, group it in biologically sensible subtypes and improve our knowledge about origin and progression of ovarian cancer.

2.2 High-Throughput Profiling in Epithelial Ovarian Cancer

In the first part of this PhD thesis, we assessed the current situation of high-throughput profiling technology biomarker studies in epithelial ovarian cancer. The aim was to see which studies have been performed from both a technical and a results point of view. A systematic online database search was performed on all published high-throughput technologies ('omics') which focused on epithelial ovarian cancer. The goal was to bring the state-of-the-art proteogenomic studies into translational context, showing the co-evolution of ovarian cancer knowledge based on the results of profiling technologies. We summarized the data and reviewed the clinical questions addressed in the published studies, highlighting the most promising novel biomarkers (1). We also described directions for future research, based on technical limitations and identified needs. Using this approach, we identified the emerging field of glycomics as a new potential for epithelial ovarian cancer biomarker research. Glycomics has not been used for this purpose yet, but is rapidly progressing into the mainstream of biology and biomedicine. The glycome represents the main class of post-translational modifications and provides access to circulating anti-glycan antibodies in a patient's blood (2-5). This observation could have two major clinical applications in epithelial ovarian cancer: (A) novel biomarker detection and (B) targets for immunotherapy.

In order to perform a discovery approach we established a biobank including prospective sample and data collection. Collection and processing of blood serum, ascites and tissue has been performed at the Department of Gynaecology (University Hospital Zurich), after written informed consent (Ethical Approval 2006, SPUK, Canton of Zurich, Switzerland)

was granted. For each sample, comprehensive clinical and experimental parameters were stored in our PEROV- database.

2.3 **Biomarker Research Using Printed Glycan Array**

Aberrant glycosylation of proteins and lipids during malignant transformation results in the appearance of specific glycan structures known as Tumor Associated Carbohydrate Antigens, or TACA, on cell surfaces and serum components (6, 7). Combinations of TACA are always present during malignant transformation. Potential cancer tumor markers which can be detected in serum of ovarian cancer patients are based on the observation that some TACA are shed into or present in the bloodstream, reflecting the pathological state. Our immune system constantly surveys the body for such aberrant carbohydrate structures and generates a response against them. Consequently, naturally occurring anti-glycan antibodies binding to these structures are promising targets for detection of epithelial ovarian cancer.

The second part of this PhD thesis we detected individual patient anti-glycan antibody signatures as well as multivariate patterns in the form of ovarian cancer-associated anti-glycan antibodies with an identified diagnostic value. Using printed glycan array technology, we screened 57 well-defined and prospectively collected serum samples of patients with either non-mucinous ovarian cancer or borderline tumor, and healthy control patients at the time of surgery.

The three essential key features we were using are: (**1**) in-house printing of custom glycan arrays, (**2**) established experimental development of printed glycan arrays and (**3**) printed glycan array-dedicated mathematical data processing tools (Figure 6). The production of standardized PGA slides for simultaneous profiling of anti-glycan antibodies involved also the development of a glycan library containing over 200 glycans, including TACA, natural oligosaccharides, blood group antigens and their related structures. The printing protocol implemented quality control of glycan arrays, such as print spot quality, batch-to-batch reproducibility and standardization of printed glycan array slides. We developed standard operating procedures for serum development and profiling of anti-glycan antibodies of immunoglobulin subtype IgM, IgG and IgA. Standardization of serum development also included bar-code and tracking systems to link printed glycan slides to individual serum samples. Each serum incubated slide was proved after fluorescence scan and low quality developments were repeated. For technical validations approximately 75% of all samples were screened in duplicates. Mathematical algorithms were developed and software implementation was used for pre-processing of printed glycan array data. Data pre-processing included

transformation and normalization of raw glycan array data to minimize technical variability and maximize performance. Data validation and statistical examination for biomarker detection as well as study of glycan-binding motifs were performed using the open source software CRAN R –project. Statistical analysis included detection of univariate and multivariate discriminative sets of glycans (biomarkers) and was further evaluated in terms of predictive precision, sensitivity and specificity. Glycan-binding motifs were studied by unsupervised learning.

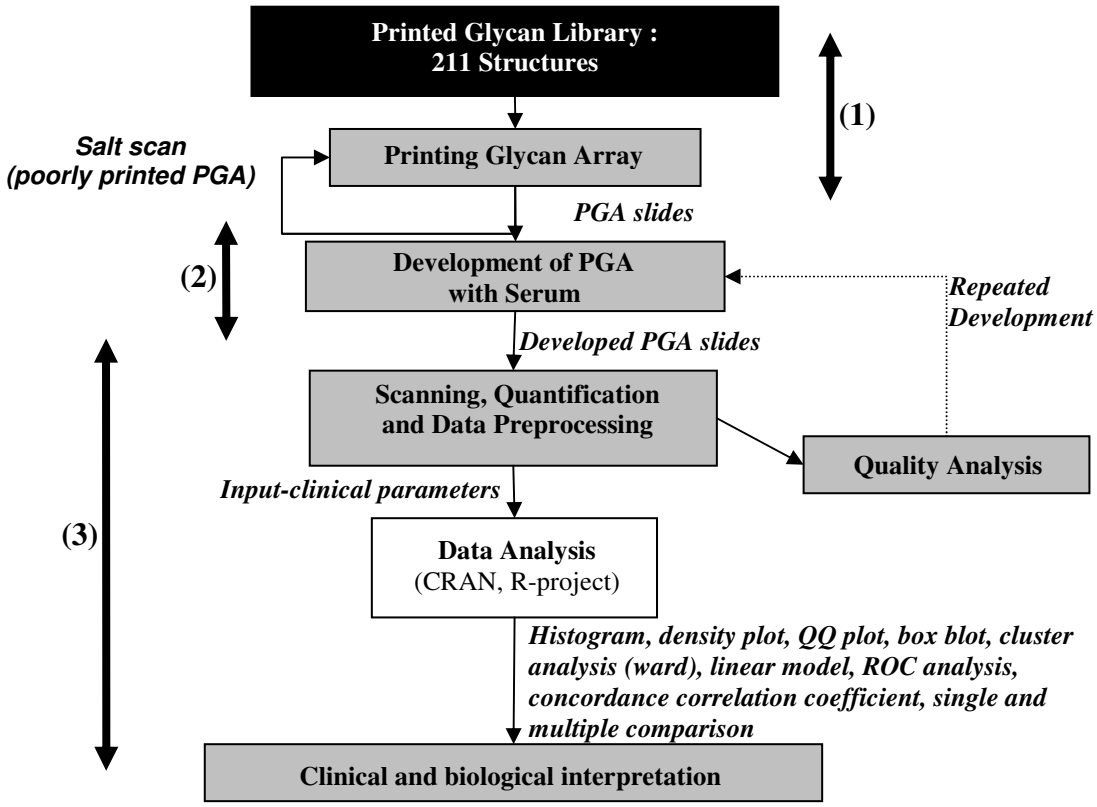


Figure 6 Process of printed glycan array experiment.
Essential key features are in-house printing of glycan arrays (**1**), serum development of printed glycan arrays (**2**), and data processing (**3**), including scanning, quantification/ data preprocessing and statistical analysis for clinical and biological interpretation.

2.4 Study of Glycan-Binding Motifs

The study of glycan-binding motifs is a major field in glycobiology which is necessary to understand molecular recognition events. A number of glycan-based immunoassays for high-throughput investigations of the glycome have been developed recently to analyze the binding between glycans and glycan-binding proteins (*e.g.* lectins). To our knowledge, no study has examined the binding motifs of anti-glycan antibodies to immobilized glycans in healthy and ovarian cancer samples. Within our study cohort we aimed to identify specific glycan-binding patterns of naturally occurring anti-glycan antibodies. We further performed a

comprehensive analysis by applying different glycan-based immunoassays to study chemically synthesized glycans displayed by different ways to human anti-glycan antibodies. In this third part of the PhD thesis we examined human serum samples for their detection rates in three glycan immunoassays: printed glycan array, custom made sandwich ELISA technique and multiplex flow cytometric suspension array. In all three assays we aimed for the detection of anti-blood group A/B antibodies using A/B trisaccharides, known to be AB0 blood group minimal determinants. Related antigens of AB0 blood group were used as biological validation technique for our new PGA approach because it is widely used in immunochemical studies and well known for biological pathways. We did also assess the detection rate of our PGA generated candidate with the highest ovarian cancer detection rate, P_1 (*Galα1-4Galβ1-4GlcNAcβ*) on all glycan-based immunoassays.

2.5 **Identification of Glycan-Binding Proteins**

Oligosaccharides presented on eukaryotic cell surfaces are known targets for anti-glycan antibodies as well as other glycan binding proteins, such as galectins or siglecs. Glycan-binding proteins specifically recognize and interact with carbohydrates, and can be found in every type of known organism. They play major roles in biological processes such as immune recognition and regulation, inflammatory responses, cytokine signaling, and cell adhesion. As fourth part of this PhD thesis, we tried to identify specific proteins that bind to a high-purity chemically synthesized oligosaccharide. We established an experimental strategy to study functionality and a complex hierarchical system of specific glycan-protein interactions. This includes affinity purification of potential glycan-binding partners from pooled ascites, validation of eluted proteins by SDS-PAGE followed by silver staining, immuno-detection of glycan-binding immunoglobulins, and validation of anti-glycan antibodies by sandwich ELISA and printed glycan array methodology. Final identification of potential glycan-binding proteins was performed by liquid chromatography mass spectrometry (LC-MS/MS).

2.6 **References**

1. Jacob F, Goldstein DR, Fink D, & Heinzelmann-Schwarz V (2009) Proteogenomic studies in epithelial ovarian cancer: established knowledge and future needs. *Biomarkers Med.* 3(6):743-756.
2. Dotan N, Altstock RT, Schwarz M, & Dukler A (2006) Anti-glycan antibodies as biomarkers for diagnosis and prognosis. *Lupus* 15(7):442-450.
3. Huflejt ME*, et al.* (2009) Anti-carbohydrate antibodies of normal sera: findings, surprises and challenges. *Mol Immunol* 46(15):3037-3049.
4. Oyelaran O, McShane LM, Dodd L, & Gildersleeve JC (2009) Profiling Human Serum Antibodies with a Carbohydrate Antigen Microarray. *J Proteome Res*.

5. Wandall HH, *et al.* (2010) Cancer biomarkers defined by autoantibody signatures to aberrant O-glycopeptide epitopes. *Cancer Res* 70(4):1306-1313.
6. Hakomori S (2001) Tumor-associated carbohydrate antigens defining tumor malignancy: basis for development of anti-cancer vaccines. *Adv Exp Med Biol* 491:369-402.
7. Hakomori S (2002) Glycosylation defining cancer malignancy: new wine in an old bottle. *PNAS* 99(16):10231-10233.

3 Results

3.1 Proteogenomic studies in epithelial ovarian cancer: established knowledge and future needs

Manuscript published in Biomarkers in Med (2009) 3(6), 743-756

Jacob F.[1], Goldstein D.R.[2,3], Fink D.[4] & Heinzelmann-Schwarz V.[1]

[1] *Translational Research Group, Department of Gynecology, University Hospital Zurich, Zurich, Switzerland*

[2] *Institut de mathématiques, Ecole Polytéchnique Fédérale, Lausanne, Switzerland*

[3] *Swiss Institute of Bioinformatics, Lausanne, Switzerland*

[4] *Department of Gynecology, University Hospital Zurich, Zurich, Switzerland*

Corresponding author: *Viola Heinzelmann-Schwarz, Translational Research Group, Department of Gynecology, University Hospital Zurich, Nord I D222, Frauenklinikstrasse 10, 8091 Zurich, Switzerland; e-mail: viola@heinzelmann.ch, Telephone: +41442555374, Telefax: +41442554553*

3.1.1 Summary

There has been a concerted effort over the last decade to improve our understanding of the complex biology of ovarian cancer. A linear growth in published proteogenomic studies has addressed a variety of questions regarding its molecular pathogenesis. A number of genes have been identified by transcriptomic approaches, some of which are being investigated as putative tumor markers (*HE4, OPN, Ep-CAM, Mesothelin*), whilst others are potential targets for molecular therapeutic approaches (*VEGF, IO4, EGFR, MUC1, CLDN4, SLPI*). Proteogenomics has the potential to further change our current characterization and treatment of ovarian cancer. Additional advances will depend on integrated study designs, interdisciplinary collaborations, use of robust high-throughput platforms, as well as uniform guidelines for bio informatic analyses.

Keywords: ovarian cancer, ovary, proteomics, transcriptomics, glycomics, profiling, microarray, diagnostics, biomarkers, post-translational modifications

3.1.2 Introduction

Epithelial ovarian cancer is the fourth most common cause of cancer death in women in the Western world and the leading cause of death from gynecological malignancies [1]. The majority of women have advanced stage disease at initial diagnosis and a five-year survival of 10 – 30% [2]. The median survival after recurrence is only two years despite advances in chemotherapy and secondary debulking surgery in selected patients [1, 3, 4]. We clearly need better insights into the biology of the disease in order to detect the disease at an earlier, potentially curable stage and to develop more targeted and effective treatment strategies.

Epithelial ovarian cancers are made up of five main histological subtypes that vary in their biological behavior, response to treatment and overall prognosis [5]. However, even within specific histological subtypes there is considerable heterogeneity, particularly within the more common serous cancers [6, 7]. Although FIGO stage and grade are still used to classify all ovarian cancers, it has been recently recognized that there are two separate and quite distinct subgroups of serous cancer: Type 1, which are low grade and indolent and have a distinct molecular pathogenesis, and the more common Type 2 serous cancers, which are high grade and disseminate rapidly [8, 9]. Type 1 tumors appear to arise from the ovarian surface epithelium or inclusion cysts [10] and follow a stepwise progression through micro invasive tumors and tumors of low malignant potential. They commonly have mutations in *KRAS* and *BRAF* [9]. In contrast, Type 2 serous tumors appear to arise from the fimbrial end of the

fallopian tube [11-14] and *p53* mutations occur early. *BRCA1/2*-associated ovarian cancers in particular appear to arise from the fimbrial end of the fallopian tube and have a better response to platinum-based chemotherapy [15-17].

Despite the genetic distinction between the various ovarian cancer histological subtypes [18], they are commonly treated in the same manner: namely, with maximal cytoreductive surgery [19-22] followed by platinum- and taxane-based chemotherapy [4, 23, 24]. It would be valuable if we could use molecular expression profiles to classify ovarian tumors into distinct subtypes based on the biological behavior and to better predict individual patient outcomes based on a such a personal molecular "biosignature" [25, 26]. Good examples of this approach are given already for breast cancers and lymphomas [27, 28].

There has been a concerted effort over the last decade to identify specific molecular profiles of ovarian cancers using cDNA microarrays and oligonucleotide arrays [29, 30]. These high-throughput technologies for transcriptomics or genomic expression profiling permit the analysis of the expression of thousands of genes in one experiment examining a specific biological system [31]. The resulting patterns allow researchers to investigate important biological questions that have not been easily addressed with traditional gene expression technologies such as Northern blots, *in situ* hybridization or RNase protection assays. Following the development of large-scale profiling technologies, the first high-throughput studies examining protein expression levels were performed using serological analysis of autologous tumor antigens by recombinant cDNA expression cloning (SEREX) [32]. Within the growing spectrum of proteomic technologies, most techniques were based on two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) mainly in combination with mass spectrometry techniques such as matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF), surface-enhanced laser desorption/ionization-time of flight (SELDI-TOF), electrospray ionization (ESI) or with liquid chromatography (LC).

There has been a linear growth in the number of publications on 'biomic' studies of ovarian cancer. Since the initial studies on transcriptomics of ovarian cancer in 1999 [33-35] 237 studies were published until the end of 2007 using various –omic technologies. A vast amount of data has been amassed over the last decade, providing greater insight into the heterogeneous genetic background of ovarian cancer. This has lead to the identification of novel markers as well as new diagnostic, prognostic and therapeutic targets.

The aim of this study is to provide a comprehensive summary of the totality of published studies pertaining to functional genomics of ovarian cancer. We attempt to put the state of the art of proteogenomic studies in ovarian cancer into context, showing the co-

evolution of ovarian cancer knowledge and technologies for new assays, which provide the basis for our view of the direction of future research. We also summarize the data and review the clinical questions addressed in the published studies and highlight the most promising novel tumor markers. We briefly describe new technical possibilities [36], which have the potential both to improve the early detection and better predict prognosis of patients with epithelial ovarian cancer, and take a glimpse at the future of ovarian cancer biomarker research.

3.1.3 Identification of genomic profiling studies.

We performed an electronic search of the online databases PubMed (http://www.ncbi.nlm.nih.gov/pubmed/), Cochrane (http://www.cochrane.org/) and Medline (http://medline.cos.com/) to identify publications on ovarian cancer proteogenomic profiling during the 9 year period 1999-2007. Keywords used to search the databases were in changing combinations: “ovarian cancer”, “ovary”, “human ovarian” in combination with “mass spectrometry”, “MS”, “serial analysis of gene expression”, “SAGE”, “SEREX”, “cDNA microarray”, “DMH”, “CGH”, “array CGH”, “expression profiles”, “oligonucleotide array”, “proteomics“, “transcriptomics“, “genomics“, or “expression profiling”, “array”, “microarray” and “SNP array”. Only published studies that used human-derived cell lines or human clinical samples were included. Articles were excluded when (a) the term “microarray” was related to tissue microarrays without further mentioning expression profiling within the paper; (b) the data were from another, already published proteogenomic study; and (c) publications in languages other than English.

The identified publications were divided into three categories: transcriptomics, proteomics, and other profiling technologies. Transcriptional profiling includes cDNA and oligonucleotide microarray studies and SAGE; proteomic profiling includes studies using SEREX, 2D PAGE, MALDI-TOF, SELDI-TOF, ESI and LC; and other profiling technologies cover differential methylation hybridization (DMH), single nucleotide polymorphism (SNP) arrays and (array) comparative genomic hybridization (aCGH) studies. The publication data are summarized according to year of publication, high-throughput technology used, number of profiled samples, sample type, sample preparation, aim of study, and main study findings. `Numbers of studies’ as listed in the supplementary data do not reflect the numbers of published studies but the research questions addressed within studies, with multiple questions possibly arising from individual proteogenomic datasets.

Data summaries were further descriptively analyzed for the overall time of 9 years and for three individual time periods of 3 years each. We evaluated the whole time period in order

to study the trend of profiling studies performed per year and method used (Figure 7). Individual 3 year periods were used to demonstrate trends within sample types (Figure 8), high-throughput technologies (Figure 9), and research questions addressed (Figure 10). Individual research questions were also further sub-classified into thirteen main categories: (a) tumor marker discovery (*Cancer vs. Normal*), (b) treatment effects (*Treatment; wild type vs. knock out or treated vs. untreated*), resistance (*Drug Resistance*), patient outcome (*Outcome*), (c) distinction of different histological subtypes (*Subtypes*), histological grades (*Grades*), the International Federation of Gynecology and Obstetrics (FIGO) stage (*Stages*), primary *versus* metastatic site (*Metastasis*), ovarian *vs.* non-ovarian cancers (*EOC vs. non-EOC*), (d) aberrant chromosomal patterns (*Chromosomes*), epigenetics (*Epigenetics*), hereditary diseases (*Hereditary*), and (e) specificity of technique (*Technical Optimization*) (Figure 10). Ovarian cancers subtypes as defined by histology, stage, grade, organ of origin and outcome were further studied in order to assess whether a genetic distinction between groups would be justified. Taken together, Figures 7- 10 trace the history of clinical biomarker identification using genomic and proteomic technologies. These figures also serve to emphasize the point that although the amount of data is increasing, we still face the daunting challenge to translate the information contained therein to clinically relevant knowledge that will benefit patients.

We incorporated published lists of new candidate tumor markers derived from transcriptomic approaches for distinguishing between healthy individuals and patients with ovarian cancer into our previously described automated over-lapping ovarian (OLOV) database [37]. All published candidate gene lists were screened for gene overlaps across individual studies, with overlaps identified by matching gene symbols, accession numbers and UniGene identifiers. We used Gene Cards (http://www.genecards.org) and the Human Protein Reference database (http://www.hprd.org) to further enhance the list of most frequently identified potential tumor marker genes, with the addition of information regarding cellular localization, translational modifications, molecular class, glycosylation as post-translational modification, available treatment options and clinical testing status (Phase I-III trials). The numbers of studies demonstrating a significant difference between research question groups are presented graphically using OLOV as described previously [37]. OLOV-incorporated genes were ranked by previously described gene identifiers and sorted in decreasing order by number of reoccurrence across studies.

3.1.4 Development of high-throughput profiling studies

Between the years 1999-2007, 237 primary studies addressing functional 'biomics' in ovarian cancer have been published. During the first five years, there is a roughly linear growth in number of publications using profiling techniques, after which the number of publications reaches a plateau (Figure 7). Two thirds of studies between 1999 and 2004 used transcriptomic (rather than proteomic or other) methods. After a clear increase in 2005 a steady decline in number of transcriptomic studies can be observed, which is linked to a nearly linear increase in proteomic and other methods since 2004. These other methods were used as frequently as proteomics methods. Although MIAME (minimal information about a microarray experiment) [38, 39] guidelines state the need to openly distribute microarray raw data for further validation or meta-analysis, most frequently only lists of significant gene expressions are published, especially for the earlier time periods. Unfortunately, such lists are of limited utility for integrating information across studies.

3.1.5 Profiling technologies used in ovarian cancer

We have grouped profiling technologies into three major classes: (1) transcriptomics (63.7% of studies): *e.g.* cDNA and oligonucleotide microarrays and SAGE technology, used to study gene expression; (2) proteomics (20.3% of studies): 2D PAGE, typically in conjunction with mass spectrometry technologies, SEREX, and protein microarray; and (3) other technologies (16.0% of studies), including: (array) CGH, SNP array and DMH. Multiple techniques were used in 15.6% of studies, most of which combine either transcriptomics with CGH (29.7%), or proteomic technologies with transcriptomics (5%) (Figure 9).

The majority (63.7%) of ovarian cancer-specific questions were addressed using technologies to identify gene expression patterns, including cDNA microarrays (51.3%) and oligonucleotide (46.3%) arrays, most commonly produced by Affymetrix (Santa Clara, California, USA; HuGeFL, U95-, U133-series).

Robust transcriptomic technologies and analysis methodologies have led to the detection of several new ovarian cancer biomarkers. In contrast, even with advances in mass spectrometry, proteomic approaches have led to only a few protein biomarkers which are used clinically. Despite technical complexity, various recent reviews have discussed the great potential of proteomics in clinical biomarker discovery [40-44]. Proteomic approaches applied to detecting clinical biomarkers typically involve a broader spectrum of samples and larger chemical information content on the protein level than other types of biomic studies [45], but technical challenges have hindered development. Due to the broad and complex diversity of proteins, the

field of proteomics has to deal with a wide range of differences in molecular size, relative abundance, concentration, polarization and ionization as well as the high dynamic concentration range of proteins in complex samples like tissues and body fluids. Since there is not at present a single, optimal instrument that satisfies all proteomics research needs, there is a wide variety of mass spectrometry instrumentations, including an extensive assortment of modular arrangements of different types of mass analysers [46]. The current standards in bioinformatical analyses [47] and normalization procedures [48] are rather limited.

Recent advances in mass spectrometry-based quantitative proteomics permit the identification of differentially expressed proteins in cells, tissues and body fluids in a high-throughput manner [49, 50]. Quantification of proteins by mass spectrometry is currently most commonly performed by differential isotopic labeling of samples, separation of the peptides by liquid chromatography and identification as well as quantification of the peptides by shotgun mass spectrometry and database searches [51-54]. Major improvements during recent years have been achieved by (a) label-free quantifications that integrate the total ion current of peptide signals as quantitative measure of original peptide concentration [55-58], (b) selected reaction monitoring for quantitative protein identification [59-61], and the development of (c) protein microarrays as protein-detecting microarrays, sandwich immunoassays with spotted antibodies, antigen capture immunoassays and direct immunoassays [62].

3.1.6 Variation of biological sample types

The most commonly profiled samples in transcriptomic studies consisted of human tissue (48.3%), cell lines (41.7%) or both tissues and cell lines (9.3%), with a small minority (0.7%) using body fluids. During recent years a change in sample selection has taken place; individual samples were required to be more uniform so that tumors with mixed histological appearances would tend to be excluded. This constraint facilitates laser capture micro dissection (LMD) [63] for paraffin-embedded or fresh-frozen tissue specimens in cancer as well as for normal controls [64], *e.g.* particularly ovarian surface epithelium. However, the rate of studies incorporating LMD in order to allow homogenous messenger RNA generation was only 8.9%. Proteomic experiments were performed on tissue specimens (14.6%), cell lines (31.2%), and body fluids, *e.g.* blood serum, ascites or pleural effusions (54.2%). The first studies examining ascites and ovarian cyst fluid have only just recently been published [65].

The selection of normal control tissue is a major concern in most of the studies published. Whilst in earlier studies whole ovaries were used as a normal control or human ovarian surface epithelial cell lines were used in comparison to tissue samples, improvements

have been under way using laser capture microdissection for human ovarian surface epithelium. Whilst quantities of ovarian surface epithelial cells are a major problem, new proposed precursor lesions like inclusion cysts and, more recently, the tubal epithelium [10-12, 14, 66] have so far only rarely been incorporated into biomic studies.

3.1.7 Study aims within genomic profiling

Study aims for all 237 publications were characterized and grouped, and are shown as absolute numbers in a bar graph (Figure 10). As study aims have changed during the years, they are presented in three-yearly intervals (1999-2001, 2002-2004, 2005-2007), thus demonstrating the general interest shift over time. During these three time periods, most studies investigated the differentiation between *cancer vs. normal* in order to find new biomarkers (82 studies), experimental *treatment* settings (48 studies), histological *subtypes* (44 studies), tumor *grades* (20 studies) and *chromosomes* (20 studies) aberrant in ovarian cancer (Figure 10). The search for a new tumor marker in ovarian cancer (*cancer vs. normal)* was the most studied aspect during 2005-2007 in 38 studies, nearly the same as during the previous three year period (2002-2004, 30 studies). The other most commonly addressed topics during the most recent three years included *treatment* effects (33 studies), *drug resistance* (31 studies), histological *subtype* distinction (28 studies), *technical optimization* (23 studies), *grades* (15 studies) and *chromosomes* (13 studies). There is a growing diversity during this most recent period, with the emergence of new fields such as *technical optimization* of experimental procedures.

3.1.8 Distinct ovarian cancer biology identified by high-throughput profiling

Clinico-pathological characteristics of ovarian cancer were examined on the molecular level to define distinct proteogenomic patterns. Studies aimed to identify a pattern for histological subtypes, stage, grade, organ of origin and disease progression. Distinctions identified for particular subgroups of ovarian tumors are shown in Figure 11. These results show that there is a clear molecular difference between histological subtypes of mucinous, clear cell and endometrioid ovarian cancers. Further, endometrioid cancers of either ovarian or endometrial origin seem to have a distinct molecular pattern, even when expressing the same histological phenotype. Serous ovarian cancers of different grades, especially serous borderline tumors, are biologically different from serous ovarian cancers, which have been further molecularly distinguished as early (FIGO stage I/II) or late (FIGO stage III/IV) stage disease.

3.1.9 Aberrantly expressed genes identified across multiple transcriptomic studies

Candidate genes differentially expressed in ovarian cancer compared to normal controls, detected by all transcriptomic studies independent of specific research question, are listed in order of their number of overlaps across studies (Table 3). Most transcriptomic profiling studies identified *OI4*, *VEGF* and *TFCP2*, from which especially *OI4* and *VEGF* have interesting implications. Both gene products are extracellulary localized, are glycosylated and are used as targets for therapies which are either already clinically used (Imatinib for OI4*;* Bevacizumab for VEGF*;* Gefitinib for EGFR) or in clinical phase II/III trials, targeting VEGF, EGFR, CLDN4 and HER3. Other less frequently identified biomolecules show involvement of either cellular communication, adhesion or immune response. Post-translational modifications, especially glycosylation modifications, can be found in 55% of these biomolecules.

We have also identified overexpressed genes overlapping across studies addressing the same specific research question. Our approach was only possible for analyses (1) Carcinoma vs. normal, (2) Subtypes, (3) Grades and (7) Treatment. In this sub question examination, a maximum of 5 overlaps across studies could be identified for individual genes (Table 4). A reason for this low number is the diversity of clinical questions which makes it difficult to give question-specific markers. Interestingly, there are some genes which persist across a number of different specific aims, namely HE4 for Carcinoma vs. normal and Grades, and OPN for "Carcinoma *vs.* normal" and "Subtypes". Genes identified in the "Treatment" research question analysis were unique to this question only (Table 4).

3.1.10 Established knowledge due to proteogenomics

The aim to improve upon early screening techniques and to develop more effective, targeted treatments for ovarian cancer has driven the field of high-throughput technologies since their inception 10 years ago. The success of the intensive international efforts to improve outcomes, with new targeted therapies being tested in phase III trials (e.g. VEGF inhibitors as the most advanced therapeutics), has now become evident. Moreover, the first diagnostic marker, HE4, is commercially promoted by Fujirebbio Inc. and is increasingly being used in conjunction with CA125. Particularly notable today is the appreciation of the heterogeneity of ovarian cancer, necessitating its re-categorisation into multiple distinct entities. It is thus recognized that we must treat especially mucinous and clear cell ovarian cancers differently from serous and endometrioid histological subtypes. The careful reflection on the employment of proteogenomic technologies and the critical assessment of the benefits gained from them compared to their costs has also become an unquestionable factor. It is equally important to

learn from mistakes and to carefully design future studies in order to maximise the benefit from these costly technologies (Table 5).

3.1.11 Future Perspectives

New, more specific high-throughput technologies under development target the biological products of the gene-protein network, namely glycosylation, phosphorylation and metabolisation. These reflect post-translational modifications of ovarian cancer and therefore carry important biological information. Most of our current clinical tumor markers are glycosylated proteins (CA125, CEA, CA19-9), 55% of all newly identified biomarkers (Table 3) show glycosylation. Technical simplification and standardisation of these new methods will improve the uniformity and comparability of results across cohorts and is expected to massively improve our knowledge of ovarian cancer.

During this last decade, biomics in the field of ovarian cancer has undergone a paradigm shift away from the single disease theory towards the acceptance of disease heterogeneity. This implies the presence of different biological mechanisms and precursor lesions for the development of ovarian cancer and the requirement for differential classifications and individualized disease subtype-specific treatments. One significant consequence is that shortcomings of previous studies, such as heterogeneity of samples and methods, application of inadequate methodologies or missing bioinformatics expertise in analysing the datasets, have been largely overcome. Another consequence is that numerous tumor markers have been proposed and are now under evaluation in combination with or without CA125, the still unbeaten tumor marker for ovarian cancer.

Advantages of previous proteogenomic studies include the wealth of data available within the scientific community, the lessons learned from mistakes, and subsequent improvements in technologies and study design and analysis techniques. A far-reaching aspect of the increased data acquisition is an accompanying increase in data sharing. There are now several prominent publicly accessible databases of genomic information. These include general microarray data repositories such as the Gene Expression Omnibus (GEO) [67, 68], ArrayExpress [69], and the Stanford Microarray Database (SMD) [70] and cancer-specific data repositories such as the Expression Project for Oncology (expO, http://www.intgen.org/expo.cfm) and ONCOMINE [71, 72].

The experimental design and statistical analysis of microarray studies has evolved greatly from the early days of simplistic, single-sample comparisons [73]. In addition to the ordinary considerations of any scientific study, important aspects more specific to design of

microarray studies include pairing of samples for dual-channel arrays and sample processing protocols to avoid confounding and increase precision of inferences [74]. The analysis of microarray data generally includes a technology-specific component (so-called 'low-level' analysis) and a study aim-specific ('high-level') analysis. Low-level analysis consists of data preprocessing (image analysis and normalization), required to reliably quantify fluorescence intensities for each transcript. For a review of methods for cDNA arrays, see Smyth and Speed 2003 [75]; for Affymetrix and other single-channel arrays, the most widely used methods are MAS 5 [76] and RMA [77, 78]. The most commonly carried out type of higher level analysis is identification of differential expression. Many methods exist for this, but the most reliable ones involve computing a type of *t*-statistic (or *F*-statistic) based on (log) fold change and a moderated standard deviation. A widely used method is the linear modeling approach [79].

As data sharing becomes the new standard, we gain increased opportunities for knowledge synthesis. Among these is the ability to carry out statistically sound meta-analysis for clinically relevant questions. Integrated analysis of high-throughput transcriptomic studies is complex. The ONCOMINE web tool (http://www.oncomine.org/) has some meta-analytic capabilities, however it provides limited flexibility. We have created a more adaptable approach for combining information across gene expression studies [80, 81] but, like ONCOMINE, it still relies on having access to primary data (not just a summary *p*-value or ranked gene list). An extension of our methodology should allow for combination of heterogeneous data types (such as gene expression and glycomic data) rather than across just gene expression, thereby making fuller use of existing data resources.

Despite widespread acceptance and requirement of MIAME guidelines [38], data sharing is still incomplete, limiting the possibility to quantitatively combine and critically reflect on all research results.

Based on this timeline examination of recent ovarian cancer research and current technological trends it is of major concern that only a limited number of genes are detectable in an overlap approach. This minimal number of overlaps can be explained by the limited number of available datasets, variation in study design and platforms used and diversity of research questions studied. For this field to be of more efficient assistance in identifying clinically usable biomarkers individual studies need to be comparable within each other. Based on our examination we suggest that future biomic studies meet the following requirements:

(A) adequate number of samples, depending on the research question preferably from different continents (in order to avoid or address population-specific effects),

(B) prospective inclusion of patients where appropriate for the study question,

(C) inclusion of all presumed precursor lesions, namely ovarian surface epithelium, inclusion cysts, fimbrial end of fallopian tube epithelium,
(D) inclusion of other cancers to distinguish ovarian specificity,
(E) absence of mixed pathological histotypes,
(F) clear-cut study question (or hypothesis) and study groups, and
(G) microdissection in all tissue samples.

Of particular importance is that

(H) qualified analysts (e.g. bioinformaticians, statisticians) perform the data analyses according to
(I) internationally accepted guidelines.

Such guidelines need to be developed to establish clear requirements for data analyses. At present, MIAME compliance is widespread but not universal. The research community will need to adapt quickly to new scientific insights in ovarian cancer to avoid costly, unnecessarily repetitive studies. To this end, we encourage the establishment and support of internationally-driven programs and open access platforms that guarantee the sharing of data, especially for the new technologies under current development.

3.1.12 Executive Summary

Epithelial ovarian cancer is the leading cause of death from gynecological malignancies

- Survival rate of 20% in advanced stage disease has not changed over the last 10 years, even with improvements in chemotherapy regimen
- Increasing acceptance of the heterogeneity of ovarian cancer, which requires a disease subtype-specific treatment regimen
- Proteogenomics will be the basis for improvements in ovarian cancer diagnosis and management in the future

High-throughput profiling technologies and their development in ovarian cancer

- 237 proteogenomic primary studies were published on ovarian cancer during 1999-2007
- Most studies focused on the detection of new clinical biomarkers for ovarian cancer
- Transcriptomic (gene expression) approaches were most commonly used
- Proteomic methods are gaining stability and reproducibility during the last few years and will have major potential in the future

Heterogeneity as the main ovarian cancer biology finding using proteogenomic high-throughput profiling

- A distinct genomic background was found for mucinous, clear cell and endometrioid ovarian cancers as compared to serous ovarian cancers
- Serous borderline tumors demonstrate a distinct genetic origin from high grade, high stage serous ovarian cancers
- The potential ovarian cancer precursor lesion has been defined as either ovarian surface epithelium, inclusion cysts, fallopian tubal epithelium, or tubal epithelium at the fimbrial end of the fallopian tube

Overexpressed genes identified across multiple transcriptomic studies

- Most transcriptomic profiling studies identified OI4, VEGF and TFCP2 as overexpressed genes in ovarian cancer
- 55% of all identified transcriptomic biomarkers express glycosylation
- Multiple identified biomarkers are already used as new targeted treatments and for diagnostic purposes

Conclusions

- Development of new high-profiling technologies targeting the biological products of the gene-protein network
- Glycomics will be an area of particular interest in the development of new biomarkers in the future as most so far identified tumor markers show glycosylation changes
- Establishment and support of internationally-driven, multidisciplinary programs joining datasets and cohorts
- Study design requirements are: question-driven prospectively collected clinical samples incorporating all proposed precursor lesions, clear histological subtypes generated using microdissection
- Definition of bioinformatical analysis standards and integration of different data types
- Development of internationally accepted guidelines for bioinformatical analyses using qualified analysts
- Open access platforms requiring full data distribution

3.1.13 References

1. Ozols, RF: Challenges for chemotherapy in ovarian cancer. *Ann Oncol* 17 Suppl 5, v181-7 (2006).
2. Fishman, DA, K Bozorgi: The scientific basis of early detection of epithelial ovarian cancer: the National Ovarian Cancer Early Detection Program (NOCEDP). *Cancer Treat Res* 107, 3-28 (2002).

3. Ozols, RF: Systemic therapy for ovarian cancer: current status and new treatments. *Semin Oncol* 33, S3-11 (2006).
4. Agarwal, R, M Linch, SB Kaye: Novel therapeutic agents in ovarian cancer. *Eur J Surg Oncol* 32, 875-86 (2006).
5. Gilks, CB: Subclassification of ovarian surface epithelial tumors based on correlation of histologic and molecular pathologic data. *Int J Gynecol Pathol* 23, 200-5 (2004).
6. Hess, V, R A'Hern, N Nasiri *et al.*: Mucinous epithelial ovarian cancer: a separate entity requiring specific treatment. *JCO* 22, 1040-1044 (2004).
7. Heinzelmann-Schwarz, VA, M Gardiner-Garden, SM Henshall *et al.*: A distinct molecular profile associated with mucinous epithelial ovarian cancer. *Br J Cancer* 94, 904-13 (2006).
8. Dehari, R, RJ Kurman, S Logani, M Shih Ie: The development of high-grade serous carcinoma from atypical proliferative (borderline) serous tumors and low-grade micropapillary serous carcinoma: a morphologic and molecular genetic analysis. *Am J Surg Pathol* 31, 1007-12 (2007).
9. Kurman, RJ, M Shih Ie: Pathogenesis of ovarian cancer: lessons from morphology and molecular biology and their clinical implications. *Int J Gynecol Pathol* 27, 151-60 (2008).
10. Auersperg, N, AS Wong, KC Choi, SK Kang, PC Leung: Ovarian surface epithelium: biology, endocrinology, and pathology. *Endocr Rev* 22, 255-88 (2001).
11. Kindelberger, DW, Y Lee, A Miron *et al.*: Intraepithelial carcinoma of the fimbria and pelvic serous carcinoma: Evidence for a causal relationship. *Am J Surg Pathol* 31, 161-9 (2007).
**12. Lee, Y, A Miron, R Drapkin *et al.*: A candidate precursor to serous carcinoma that originates in the distal fallopian tube. *J Pathol* 211, 26-35 (2007).

Identification of the fimbrial end of the fallopian tube as a precursor lesion in BRCA based serous ovarian cancers

13. Jarboe, E, A Folkins, MR Nucci *et al.*: Serous carcinogenesis in the fallopian tube: a descriptive classification. *Int J Gynecol Pathol* 27, 1-9 (2008).
14. Salvador, S, A Rempel, RA Soslow, B Gilks, D Huntsman, D Miller: Chromosomal instability in fallopian tube precursor lesions of serous carcinoma and frequent monoclonality of synchronous ovarian and fallopian tube mucosal serous carcinoma. *Gynecol Oncol* (2008).
15. Marcelis, CL, HW van der Putten, C Tops, LC Lutgens, U Moog: Chemotherapy resistant ovarian cancer in carriers of an hMSH2 mutation? *Fam Cancer* 1, 107-9 (2001).
16. Weberpals, JI, KV Clark-Knowles, BC Vanderhyden: Sporadic epithelial ovarian cancer: clinical relevance of BRCA1 inhibition in the DNA damage and repair pathway. *J Clin Oncol* 26, 3259-67 (2008).
17. Cass, I, RL Baldwin, T Varkey, R Moslehi, SA Narod, BY Karlan: Improved survival in women with BRCA-associated ovarian carcinoma. *Cancer* 97, 2187-95 (2003).
18. Lawrenson, K, SA Gayther: Ovarian cancer: a clinical challenge that needs some basic answers. *PLoS Med* 6, e25 (2009).
19. Akahira, JI, H Yoshikawa, Y Shimizu *et al.*: Prognostic factors of stage IV epithelial ovarian cancer: a multicenter retrospective study. *Gynecol Oncol* 81, 398-403 (2001).
20. Crawford, SC, PA Vasey, J Paul, A Hay, JA Davis, SB Kaye: Does aggressive surgery only benefit patients with less advanced ovarian cancer? Results from an international comparison within the SCOTROC-1 Trial. *J Clin Oncol* 23, 8802-11 (2005).
21. Hacker, NF, JS Berek, LD Lagasse, RK Nieberg, RM Elashoff: Primary cytoreductive surgery for epithelial ovarian cancer. *Obstet Gynecol* 61, 413-20 (1983).

22. Eltabbakh, GH, SL Mount, B Beatty, L Simmons-Arnold, K Cooper, A Morgan: Factors associated with cytoreducibility among women with ovarian carcinoma. *Gynecol Oncol* 95, 377-83 (2004).
23. Bhoola, S, WJ Hoskins: Diagnosis and management of epithelial ovarian cancer. *Obstet Gynecol* 107, 1399-410 (2006).
24. Friedlander, M, M Buck, D Wyld *et al.*: Phase II study of carboplatin followed by sequential gemcitabine and paclitaxel as first-line treatment for advanced ovarian cancer. *Int J Gynecol Cancer* 17, 350-8 (2007).
25. Carr, KM, K Rosenblatt, EF Petricoin, LA Liotta: Genomic and proteomic approaches for studying human cancer: prospects for true patient-tailored therapy. *Hum Genomics* 1, 134-40 (2004).
26. Yap, TA, CP Carden, SB Kaye: Beyond chemotherapy: targeted therapies in ovarian cancer. *Nat Rev Cancer* 9, 167-81 (2009).
27. Sotiriou, C, SY Neo, LM McShane *et al.*: Breast cancer classification and prognosis based on gene expression profiles from a population-based study. *PNAS* 100, 10393-10398 (2003).
28. Bhojwani, D, H Kang, RX Menezes *et al.*: Gene Expression Signatures Predictive of Early response and Outcome in High-Risk Childhood Acute Lymphoplastic Leukemia: A Children's Oncology Group Study. *J Clin Oncol* 26, 4376-84 (2008).
29. Schena, M, D Shalon, RW Davis, PO Brown: Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270, 467-70 (1995).
30. Shalon, D, SJ Smith, PO Brown: A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization. *Genome Res* 6, 639-45 (1996).
31. Afshari, CA, EF Nuwaysir, JC Barrett: Application of complementary DNA microarray technology to carcinogen identification, toxicology, and drug safety evaluation. *Cancer Res* 59, 4759-60 (1999).
32. Naora, H, FJ Montz, CY Chai, RB Roden: Aberrant expression of homeobox gene HOXA7 is associated with mullerian-like differentiation of epithelial ovarian tumors and the generation of a specific autologous antibody response. *Proc Natl Acad Sci U S A* 98, 15209-14 (2001).
33. Schummer, M, WV Ng, RE Bumgarner *et al.*: Comparative hybridization of an array of 21,500 ovarian cDNAs for the discovery of genes overexpressed in ovarian carcinomas. *Gene* 238, 375-85 (1999).
34. Duan, Z, AJ Feller, RT Penson, BA Chabner, MV Seiden: Discovery of differentially expressed genes associated with paclitaxel resistance using cDNA array technology: analysis of interleukin (IL) 6, IL-8, and monocyte chemotactic protein 1 in the paclitaxel-resistant phenotype. *Clin Cancer Res* 5, 3445-53 (1999).
35. Wang, K, L Gan, E Jeffery *et al.*: Monitoring gene expression profile changes in ovarian carcinomas using cDNA microarray. *Gene* 229, 101-8 (1999).
36. Pandey, A, M Mann: Proteomics to study genes and genomes. *Nature* 405, 837-46 (2000).
37. Heinzelmann-Schwarz, VA, M Gardiner-Garden, SM Henshall *et al.*: Overexpression of the cell adhesion molecules DDR1, Claudin 3, and Ep-CAM in metaplastic ovarian epithelium and ovarian cancer. *Clin Cancer Res* 10, 4427-36 (2004).
38. Knudsen, TB, GP Daston: MIAME guidelines. *Reprod Toxicol* 19, 263 (2005).
39. Brazma, A, P Hingamp, J Quackenbush *et al.*: Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nat Genet* 29, 365-71 (2001).

40. Annunziata, CM, N Azad, AS Dhamoon, G Whiteley, EC Kohn: Ovarian cancer in the proteomics era. *Int J Gynecol Cancer* 18 Suppl 1, 1-6 (2008).
41. Kohn, EC, N Azad, C Annunziata, AS Dhamoon, G Whiteley: Proteomics as a tool for biomarker discovery. *Dis Markers* 23, 411-7 (2007).
42. Nossov, V, M Amneus, F Su *et al.*: The early detection of ovarian cancer: from traditional methods to proteomics. Can we really do better than serum CA-125? *Am J Obstet Gynecol* 199, 215-23 (2008).
*43. Sasaroli, D, G Coukos, N Scholler: Beyond CA125: the coming of age of ovarian cancer biomarkers. *Biomarkers in Medicine* 3, 275-288 (2009).
44. Tung, CS, KK Wong, SC Mok: Biomarker discovery in ovarian cancer. *Women's Health (Lond Engl)* 4, 27-40 (2008).
45. Hanash, S: Disease proteomics. *Nature* 422, 226-32 (2003).
46. Schuchardt, S, A Sickmann: Protein identification using mass spectrometry: a method overview. *EXS* 97, 141-70 (2007).
47. Biron, DG, C Brun, T Lefevre *et al.*: The pitfalls of proteomics experiments without the correct use of bioinformatics tools. *Proteomics* 6, 5577-96 (2006).
48. Ting, L, MJ Cowley, SL Hoon, M Guilhaus, MJ Raftery, R Cavicchioli: Normalization and statistical analysis of quantitative proteomics data generated by metabolic labeling. *Mol Cell Proteomics* (2009).
49. Domon, B, R Aebersold: Mass spectrometry and protein analysis. *Science* 312, 212-7 (2006).
50. Aebersold, R,M Mann: Mass spectrometry-based proteomics. *Nature* 422, 198-207 (2003).
51. Ross, PL, YN Huang, JN Marchese *et al.*: Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics* 3, 1154-69 (2004).
52. Ong, ES, SM Len, AC Lee: Differential protein expression of the inhibitory effects of a standardized extract from Scutellariae radix in liver cancer cell lines using liquid chromatography and tandem mass spectrometry. *J Agric Food Chem* 53, 8-16 (2005).
53. Pierce, JD, M Fakhari, KV Works, JT Pierce, RL Clancy: Understanding proteomics. *Nurs Health Sci* 9, 54-60 (2007).
54. Schmidt, A, J Kellermann, F Lottspeich: A novel strategy for quantitative proteomics using isotope-coded protein labels. *Proteomics* 5, 4-15 (2005).
55. Cox, J, NC Hubner, M Mann: How much peptide sequence information is contained in ion trap tandem mass spectra? *J Am Soc Mass Spectrom* 19, 1813-20 (2008).
56. Mueller, M, FT Peters, GA Ricaurte, HH Maurer: Validated liquid chromatographic-electrospray ionization mass spectrometric assay for simultaneous determination of 3,4-methylenedioxymethamphetamine and its metabolites 3,4-methylenedioxyamphetamine, 3,4-dihydroxymethamphetamine, and 4-hydroxy-3-methoxymethamphetamine in squirrel monkey plasma. *J Chromatogr B Analyt Technol Biomed Life Sci* 855, 262-70 (2007).
57. Sturm, S, J Schinnerl, H Greger, H Stuppner: Nonaqueous capillary electrophoresis-electrospray ionization-ion trap-mass spectrometry analysis of pyrrolo- and pyrido[1,2-a]azepine alkaloids in Stemona. *Electrophoresis* 29, 2079-87 (2008).
58. Wollscheid, B, D Bausch-Fluck, C Henderson *et al.*: Mass-spectrometric identification and relative quantification of N-linked cell surface glycoproteins. *Nat Biotechnol* 27, 378-86 (2009).
59. Keshishian, H, T Addona, M Burgess, E Kuhn, SA Carr: Quantitative, multiplexed assays for low abundance proteins in plasma by targeted mass spectrometry and stable isotope dilution. *Mol Cell Proteomics* 6, 2212-29 (2007).

60. Lange, V, P Picotti, B Domon, R Aebersold: Selected reaction monitoring for quantitative proteomics: a tutorial. *Mol Syst Biol* 4, 222 (2008).
61. Whiteaker, JR, H Zhang, L Zhao *et al.*: Integrated pipeline for mass spectrometry-based discovery and confirmation of biomarkers demonstrated in a mouse model of breast cancer. *J Proteome Res* 6, 3962-75 (2007).
62. MacBeath, G: Protein microarrays and proteomics. *Nat Genet* 32 Suppl, 526-32 (2002).
63. Simone, NL, RF Bonner, JW Gillespie, MR Emmert-Buck, LA Liotta: Laser-capture microdissection: opening the microscopic frontier to molecular analysis. *Trends Genet* 14, 272-6 (1998).
64. Hergenhahn, M, M Kenzelmann, HJ Grone: Laser-controlled microdissection of tissues opens a window of new opportunities. *Pathol Res Pract* 199, 419-23 (2003).
65. Lemaire, R, SA Menguellet, J Stauber *et al.*: Specific MALDI imaging and profiling for biomarker hunting and validation: fragment of the 11S proteasome activator complex, Reg alpha fragment, is a new potential ovary cancer biomarker. *J Proteome Res* 6, 4127-34 (2007).
66. Jarboe, EA, AK Folkins, R Drapkin, TA Ince, ES Agoston, CP Crum: Tubal and ovarian pathways to pelvic epithelial cancer: a pathological perspective. *Histopathology* (2008).
67. Barrett, T, TO Suzek, DB Troup *et al.*: NCBI GEO: mining millions of expression profiles--database and tools. *Nucleic Acids Res* 33, D562-6 (2005).
68. Edgar, R, M Domrachev, AE Lash: Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* 30, 207-10 (2002).
69. Parkinson, H, M Kapushesky, N Kolesnikov *et al.*: ArrayExpress update--from an archive of functional genomics experiments to the atlas of gene expression. *Nucleic Acids Res* 37, D868-72 (2009).
70. Demeter, J, C Beauheim, J Gollub *et al.*: The Stanford Microarray Database: implementation of new analysis tools and open source release of software. *Nucleic Acids Res* 35, D766-70 (2007).
*71. Rhodes, DR, S Kalyana-Sundaram, V Mahavisno *et al.*: Oncomine 3.0: genes, pathways, and networks in a collection of 18,000 cancer gene expression profiles. *Neoplasia* 9, 166-80 (2007).

Updates the contents and functionality of Oncomine

*72. Rhodes, DR, J Yu, K Shanker *et al.*: ONCOMINE: a cancer microarray database and integrated data-mining platform. *Neoplasia* 6, 1-6 (2004).

Describes the contents and functionality of Oncomine

73. Goldstein, DR, M Delorenzi, *Statistical design and Data Analysis for Microarray Experiments*, in *Unraveling Lipid Metabolism with Microarrays*, A Berger and MA Roberts (eds.), Editor. 2004, Dekker: New York.
74. Yang, YH, T Speed: Design issues for cDNA microarray experiments. *Nat Rev Genet* 3, 579-88 (2002).
75. Smyth, GK, T Speed: Normalization of cDNA microarray data. *Methods* 31, 265-73 (2003).
76. *Affymetrix Microarray Suite User Guide*. 2002, Affymetrix: Santa Clara, CA.
77. Irizarry, RA, BM Bolstad, F Collin, LM Cope, B Hobbs, TP Speed: Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res* 31, e15 (2003).
78. Irizarry, RA, B Hobbs, F Collin *et al.*: Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 4, 249-64 (2003).
79. Smyth, GK: Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 3, Article3 (2004).

**80. Wirapati, P, DR Goldstein, M Delorenzi, *Integrated Analysis of Gene Expression Profiling Studies - Examples in Breast Cancer*, in *Bioinformatics a Swiss Perspective*, R.D. Appel and E. Feytmans, Editors. 2009, World Scientific Pub Co.
Details a pipeline for meta-analysis of microarray data, and illustrates the methods with two breast cancer examples
**81. Wirapati, P, C Sotiriou, S Kunkel *et al.*: Meta-analysis of gene expression profiles in breast cancer: toward a unified understanding of breast cancer subtyping and prognosis signatures. *Breast Cancer Res* 10, R65 (2008).
An example of meta-analysis of gene expression profiles in breast cancer with both the analytic and biological perspectives

3.1.14 Figures

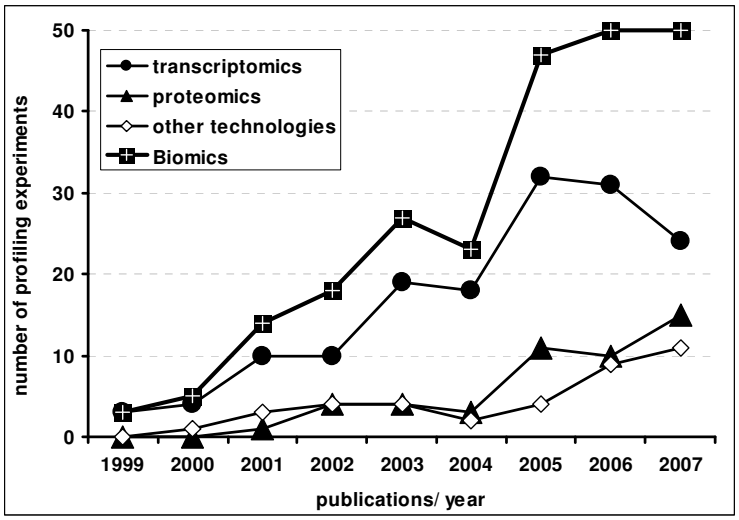


Figure 7 Development of biomics in ovarian cancer from 1999-2007.
Absolute numbers of profiling experiments published over the last 9 years. All incorporated techniques are grouped within the "Biomics" graph and individually for "transcriptomics" (cDNA microarray, oligonucleotide array and SAGE), "proteomics" (SEREX, 2D PAGE, protein microarray, SELDI -, ESI- and MALDI-TOF MS and their modified techniques) and "other technologies" (SAGE, SNP array, CGH and DMH).

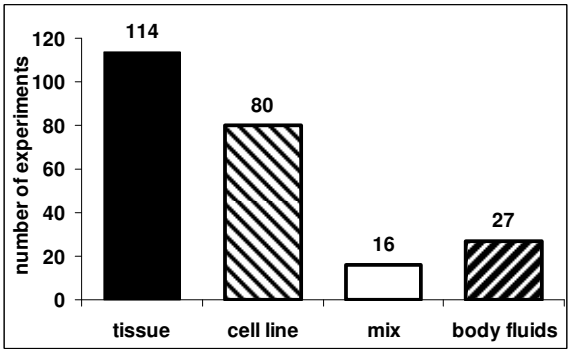


Figure 8 Variation of profiled biological samples using high-throughput technologies.
Samples used for profiling experiments analysed within this study: tissue samples ("tissue", n=114), ovarian cancer cell lines ("cell line", n=80), and a mix of both ("mix", n=16). "Body fluids" (n=27) stands for biological samples from blood, pleural effusions or ascites.

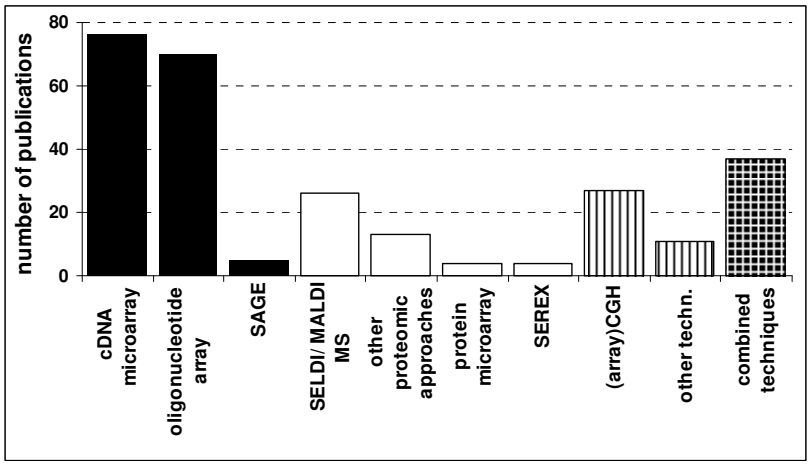


Figure 9 Ovarian cancer profiling studies using high-throughput technology.
Numbers of publications using specific high-throughput technologies are shown in bar graphs. Technologies are grouped into transcriptomics (black bars), proteomics (white bars), and other technologies of proteogenomics (shaded graphs). "Combined techniques" represents the publications using more than one high-throughput technique.

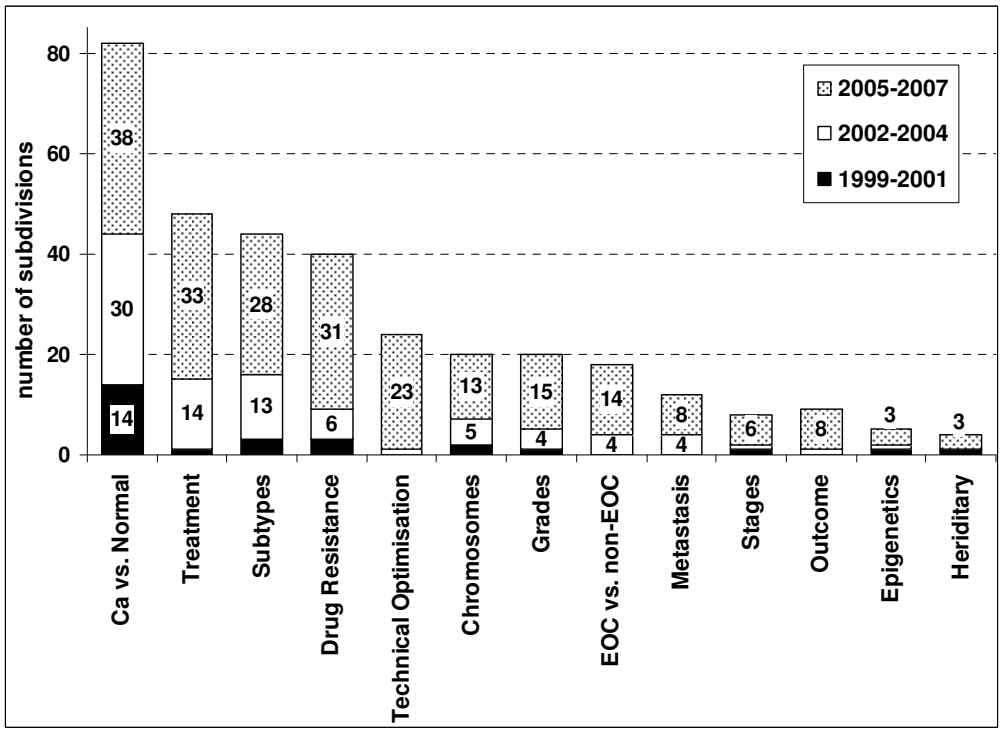


Figure 10 Categorization of investigated aims in EOC proteogenomic studies from 1999-2007.
Research questions addressed within proteogenomics studies on ovarian cancer are grouped within three-yearly time periods as (a) tumor marker discovery (*Ca vs. Normal*), (b) treatment effects (*Treatment; wild type vs. knock out or treated vs. untreated*), resistance (*Drug Resistance*), patient outcome (*Outcome*), (c) distinction of different histological subtypes (*Subtypes*), grades (*Grades*), stages (*Stages*), primary versus metastatic site (*Metastasis*), ovarian vs. non-ovarian cancers (*EOC vs. non-EOC*), (d) aberrant chromosome patterns (*Chromosomes*), epigenetics (*Epigenetics*), hereditary diseases (*Hereditary*), and (e) specificity of technique (*Technical Optimization*). Numbers within graphs represent numbers of studies within time period and question addressed.

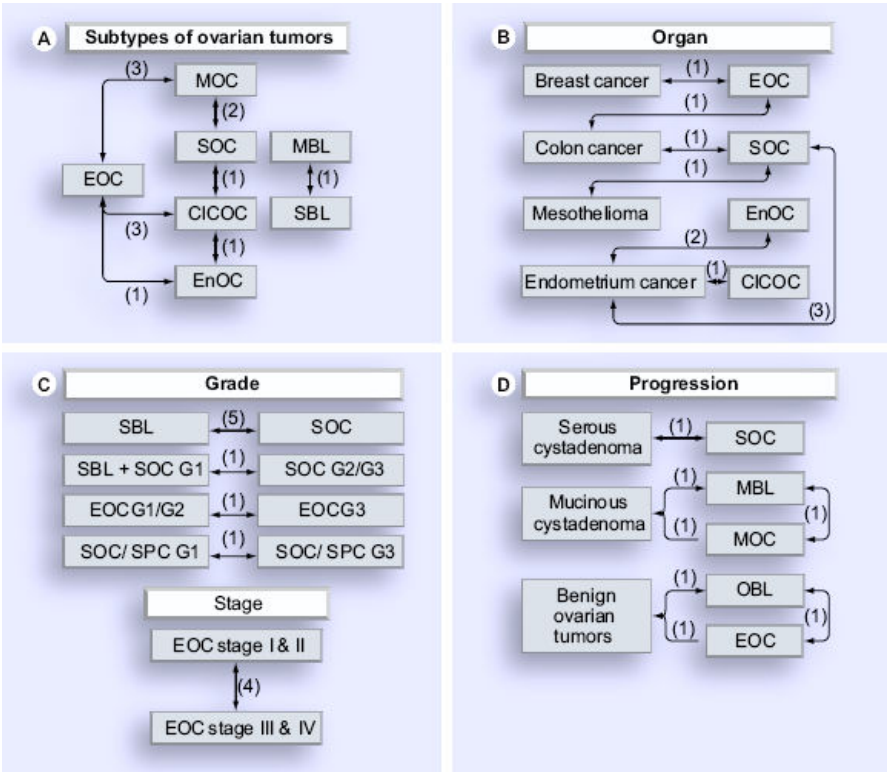


Figure 11 Genomic profiling in ovarian cancer demonstrating distinct entities.
Number of genomic profiling studies showing a genetic difference within groups of diseases studied (arrow with number of studies expressing a difference in brackets). Analyses are performed for (A) histological subtypes of ovarian tumors, (B) organ of disease origin, (C), grade and stage of ovarian cancer, (D) benign and malignant tumors within possible progression models of ovarian cancer. EOC, epithelial ovarian cancer; MOC, mucinous ovarian cancer; SOC, serous ovarian cancer; ClCOC, clear cell ovarian cancer; EnOC, endometrioid ovarian cancer; MBL, mucinous borderline tumor; SBL, serous borderline tumor.

3.1.15 Tables

Table 1. Overlapping genes most commonly identified in functional transcriptomic studies of ovarian cancer.

OLOV	Gene	Localization	PTMs	Glycan modifications	Treatment	Mechanism of (potential) treatment	Phase
11	*OI4*	Extracellular	Hydroxylation, phosphorylation, glycosylation	*O*-glycosylated			I
10	*VEGF*	Extracellular	Glycosylation	Glycosylated	Bevacizumab (Avastin®)	Monoclonal antibody to VEGF	II and III
10	*TFCP2*	Nucleus	Unknown	Unknown			
9	*BMP7*	Extracellular	Disulfide bridge	Unknown	Osigraft®	Regeneration of bone tissue	
9	*MUC1**	Plasma membrane	Glycosylation, phosphorylation, proteolytic cleavage	Heavily glycosylated		Immunogenic activity	I
9	*HE4*	Extracellular	Glycosylation	*N*-glycosylated			
8	*SLPI**	Extracellular	Proteolytic cleavage	Nonglycosylated	SLPI promoter		I
8	*POSTN*	Extracellular	Phosphorylation	Unknown		PI3K pathway	
8	*TOP2A*	Nucleus	Phosphorylation	Unknown	Etoposide phosphate (Eposin®, Vepesid®)	Topoisomerase II inhibitor	
8	*OPN*	Extracellular	Phosphorylation, glycosylation	*N*- and *O*-glycosylated		Peroxisome proliferation-activated receptors	
8	*EGFR**	Plasma membrane	Disulfide bridge, phosphorylation, glycosylation	*N*-glycosylated	Gefitinib (Iressa®)	Selective inhibitor of EGFR tyrosine kinase domain	II and III
7	*CLDN4**	Plasma membrane	Unknown	Unknown			II
7	*CDH6*	Plasma membrane	Glycosylation	Glycosylated			
7	*ALDH1*	Unknown	Acetylation	Unknown			
7	*IL6*	Extracellular	Disulfide bridge, glycosylation, phosphorylation	*N*- and *O*-glycosylated	Tocilizumab (Actemra®)	Monoclonal antibody to inhibit IL6 receptor in rheumatic diseases	
7	*Ep-CAM**	plasma membrane	Disulfide bridge, glycosylation	*N*-glycosylated	Adecatumumab (MT201)	Monoclonal antibody target tumor cells	Prostate and breast cancer clinical trials
7	*HER3**	Extracellular	Phosphorylation	Unknown	Pertuzumab, gefitinib	Selective inhibitor of EGFR tyrosine kinase domain	II and III
7	*KLK6*	Cytoplasm	Disulfide bridge, glycosylation	*N*-glycosylated			
7	*PRAME*	N/A	Unknown	Unknown			
7	*UBE2C*	Nucleus	Unknown	Unknown			
7	*p25*	Extracellular	Glycosylation	*N*-glycosylated			
7	*CD24*	Plasma membrane	Glycosylation	*N*- and *O*-glycosylated			

Most commonly identified overexpressed ovarian cancer genes using transcriptomics. Studies in which genes are overlapping, gene symbol (gene), location of analogous protein (localization), post-translational modifications (PTMs), glycan modifications, treatment and treatment options (mechanism of potential treatment) and trial status are listed.
**Genes that are already in clinical trials.*
PTM: Post-translational modification; SLPI: Secretory leukoprotease inhibitor.

Table 3 Overlapping genes most commonly identified in functional transcriptomic studies of ovarian cancer.
Most commonly identified overexpressed ovarian cancer genes (OLOV) using transcriptomics. Studies in which genes are overlapping, gene symbol (gene), location of analogous protein (localization), post-translational modifications (PTMs), glycan modifications, treatment and treatment options (mechanism of potential treatment) and trial status are listed. Genes which are already in clinical trials are underlined in grey.

Research aim	Top candidates and number of overlaps in studies
1. Carcinoma vs. Normal *max. overlaps 5*	*POSTN* (5); *HE4* (5); *OPN* (5); *Ep-CAM* (5); *ERBB3* (5); *CD24* (5); *MUC1* (5); *UBE2C* (5); *CPLA2* (5); *KLK7* (5); *PRAME* (5)
2. Subtypes *max. overlaps 4*	*OPN* (4); *POSTN* (4); *C7* (4); *CR* (4)
3. Grades *max. overlaps 3*	*HE4* (3); *IBP5* (3); *Col6A1* (3); *PAX8* (3), *ALP* (3)
7. Treatment *max. overlaps 3*	*CCN1* (4); *CYR61* (4); *GIG1* (4)

Table 4 Overlapping overexpressed genes per individual research question.

OPTIMI-ZATION	**STUDY DESIGN**	**SAMPLE SELECTION**	**BIOMICS PLATFORM SELECTION**	**BIO-INFORMATICAL ANALYSIS**
1	Experimental design in collaboration with statistician	Histological review by a specialized gynaecological pathologist	Established /robust technique	Full datasets available for analysis
2	Interdisciplinary team approach involves: biologist, gynaecological oncologist, statistician, bioinformatician, pathologist	Avoidance of mixed pathologies (histological subtypes or grades)	Technique able to address the study question	Analysis by a specialised bioinformatician
3	Prospective collection of study cohort	Microdissection of tissue (preferably by a gynaecological pathologist)	Reproducible results	Translational communication throughout complete process
4	Result validation within independent cohort (different continent)	Selection of adequate controls in tissue and serum	No inter-personal variation	Distribution of full datasets after analysis
5	Inclusion of adequate statistically significant numbers	No group combination of proven distinct entities		Inter-study validations and meta-analyses
6	Definition of study question reflects the selection of profiling technique			

Table 5 Optimization for future profiling studies.

3.2 Serum anti-glycan antibody detection of non-mucinous ovarian cancers using printed glycan array

Manuscript in preparation

Jacob F.[1], Goldstein D.R.[2,3], Bovin N.V.[4], Pochechueva T.[1], Spengler M.[1], Caduff R.[5], Fink D.[6], Vuskovic M.I.[7], Huflejt M.E.[8], Viola Heinzelmann-Schwarz[1]

[1] *Translational Research Group, University Hospital Zurich, Switzerland*

[2] *Institut de Mathématiques, Ecole Polytéchnique Fédérale, Lausanne, Switzerland*

[3] *Swiss Institute of Bioinformatics, Lausanne, Switzerland*

[4] *Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow 117997, Russian Federation*

[5] *Institute of Clinical Pathology, University Hospital Zurich, Switzerland*

[6] *Department of Gynecology, University Hospital Zurich, Switzerland*

[7] *Department of Computer Sciences, San Diego State University, San Diego, California, USA*

[8] *Cellexicon, Inc, La Jolla, CA, USA; Current address: New York University School of Medicine, Department of Cardiothoracic Surgery, Division of Thoracic Surgery and Thoracic Oncology, New York, NY, USA*

Correspondence to Viola Heinzelmann-Schwarz, Translational Research Group, Department of Gynecology, University Hospital Zurich, Nord I D222, Frauenklinikstrasse 10, 8091 Zurich, Switzerland; e-mail: viola@heinzelmann.ch, Telephone: +41442555374, Telefax: +41442554553

3.2.1 Abstract

Ovarian cancer has the highest mortality rate among gynaecological cancers. Altered glycosylation of proteins and lipids is associated with oncogenic transformation producing tumor-associated carbohydrate antigens. We investigated the potential of anti-glycan antibodies in the diagnosis of ovarian cancers using a printed glycan array containing 203 glycans. Serum samples were collected from healthy controls (n=24) and non-mucinous borderline/ovarian cancer patients of various FIGO stages (n=33). Anti-glycan antibodies bound to printed glycans were detected *via* biotin-streptavidin fluorescence. Data were validated measuring blood group associated di-, tri and tetra- saccharide antigens on known AB0 blood groups. High reproducibility in measuring anti-glycan antibodies was found overall with cluster analysis demonstrating repetitive patterns of specific core carbohydrate structures within 21 clusters: *N*-linked (n=11) *O*-linked (n=3), glycosphingolipids (n=2), not clearly sub-specifiable structures. Binary classifiers revealed 24 glycans including P_1 (*Galα1-4Galβ1-4GlcNAcβ*; $p<0.001$) significantly discriminating between ovarian cancers, borderline tumors and healthy controls. Higher sensitivity and specificity than CA125 was achieved by a panel of multivariate selected and linear combined anti-glycan antibody signals (83.3% and 84.8%, respectively). Using anti-glycan antibody profiles we detected malignant tumors with a higher sensitivity and specificity than CA125, indicating a potential for the development of a new generation of biomarkers for ovarian cancer.

Keywords: Glycan array, microarray, ovarian, cancer, anti-glycan antibodies, glycan, biomarkers

3.2.2 Introduction

Epithelial ovarian cancer is the fourth most common cause of cancer death in women in the Western world and the leading cause of death from gynecological malignancies[1]. This phenomenon is caused by the delay of initial diagnosis due to non-specific symptoms and non existent screening methods, leading to 75% of patients being detected in late stage disease, with a five-year survival of 10-20%[2]. During the last ten years a large amount of research has focused on transcriptomics and proteomics to detect a more sensitive and specific tumor marker than CA125[3,4]. Although this has advanced the biological understanding of the heterogeneity of ovarian cancer and has detected various potential biomarkers, so far, none of these seems to overcome the poor sensitivity and specificity of CA125.

Glycans are essential partners in many biological recognition processes, and the characterization of the "glycome" of cells, tissues and organisms has become one of the frontiers in the post-genomic era. The immense biological potential of altered glycosylation has been demonstrated during malignant transformation by the occurrence of tumor-associated carbohydrate antigens[5,6] as well as in diseases like rheumatoid arthritis[7] and Crohn's disease[8]. Cancer associated carbohydrates are mostly located on the surface of cancer cells and are therefore potential targets for new diagnostic biomarkers[9]. Beside their oncogenic role, naturally occurring antibodies seem to play an important part in anti-tumor surveillance, probably by binding to the repetitive motif of carbohydrate epitopes[10,11]. Anti-glycan antibodies in their malignancy-defining role identified by glycan array technology has been first described in breast cancer[12,13] and Hodgkin's lymphoma[14].

Carbohydrate arrays promise to be usable and powerful tools as they offer the potential to profile hundreds of glycan structures simultaneously. This diverse repertoire of oligosaccharides immobilized on solid matrices has the potential to map out glycan-protein interactions in a high-throughput manner[15]. Carbohydrate arrays allow determination of the specificities of glycan binding proteins[16,17], examination of microbiologically relevant glycans[18], study of carbohydrate-processing enzymes[19], and anti-glycan immune responses[20,21]. Most recent carbohydrate array studies have focused on the development of solid platforms validated by individual proteins, *e.g.* lectins or monoclonal antibodies. Microarray-based analysis of glycan–binding proteins in complex biological fluids are complicated but of highest biological relevance. To date only few publications study a small cohort of glycans and glycoconjugates within a specific biological system[22], screening anti-glycan antibodies within human serum for diagnostic purposes[20,23,24] or detecting the immune response to bacterial pathogens[25,26].

In this study we used a recently described, standardized printed glycan array to characterize discriminating carbohydrate antibody recognition in 57 human blood serum samples from healthy controls, ovarian borderline tumors and non-mucinous ovarian cancer patients of various FIGO stages. The library of printed glycans[20] contains a large number of chemically synthesized carbohydrate structures of high purity, *e.g.* blood group antigens, pathogen related oligosaccharides, lactosamines, sulphated-, sialylated-, fucosylated-carbohydrates and known as well as expected tumor associated carbohydrate antigens.

As the field of glycan-arrays is just evolving, to our knowledge there are currently only three publications describing anti-glycan antibody binding towards glycans under healthy[20,24,27], and two publication under cancer conditions, hereby detecting anti-glycan

antibodies to truncated versions of globo-H in breast cancer serum samples[28]. No anti-glycan antibody microarray profiling for the detection of ovarian cancer biomarkers has been published so far.

Application of glycan array enables the study of biomarkers in a specific disease at the level of post-translational modifications and therefore displays potentially in magnitude the presenting features of the individual state of disease[29]. Thus, this approach is capable of opening access to a broad range of novel biomarkers for ovarian cancer for diagnostic as well as therapeutic purposes. The monitoring of changes in the profiles of specific antibodies enables also the observation of the immune response of an organism against a specific disease at a specific point in time.

3.2.3 Experimental Section

Clinical cohort. Serum samples were prospectively collected from 57 women at the Department of Gynaecology University Hospital Zurich after written informed consent (Ethical approval 2006, SPUK, Canton of Zurich, Switzerland). Serum samples were collected at primary diagnosis immediately prior surgery from patients with a suspicious ovarian mass but subsequently proven negative intra-operative status (n=24) and from non-mucinous epithelial ovarian borderline tumor and cancer (EOC) patients of various FIGO stage (n=33) (Table 6). Patients with a past history of cancer or with chronic infectious diseases of non-gynaecological and gynaecological origin were excluded from the study. Routine imaging and CA125 serum tumor marker measurements (standardized ELISA, Fujirebbio Inc, Sweden) were obtained for all patients within the clinical setting. Histopathological diagnosis was independently determined by a second pathologist specialized in ovarian cancer (R.C.), with inconsistent diagnoses, mixed or mucinous histotypes being excluded. Although age-matched control patients were envisaged, the median age of women within the healthy control cohort was ten years younger than the ovarian cancer cohort (54 *versus* 63 years) (Table 6). Venous blood samples (12ml) were collected per patient in EDTA blood tubes (BD Vacutainer®, 0.184M EDTA, BD Diagnostics, Franklin Lakes, US) and stored on ice for a maximum of three hours until further processing. Blood samples were centrifuged at 3000*xg* at 4°C for 10 minutes, and aliquots of the supernatant plasma frozen at –80 °C. Comprehensive present and past medical and gynaecological history data, diagnostic imaging and serum marker measurements, surgical data including pathological results, grading, staging and residual disease were collected per patient and stored in a database.

Printed glycan array. Printed glycan array slide fabrication and high-throughput profiling was performed as previously described[20,30]. Glycans were diluted in 300mM phosphate buffer pH 8.5 containing 0.005% Tween 20 and printed by robotic pin deposition on *N*-hydroxysuccinimide activated glass slides (Nexterion Slide H, Schott, Jena, Germany). The entire glycan library containing 203 structures of 95-98% purity (Lectinity Holdings, Moscow, Russia) was printed in 50µM concentration in eight replicates. To create images documenting the deposition of each feature, printed glycan library slides were scanned for salt deposition using a ProScanArray HT Microarray Scanner (PerkinElmer) with the red reflect scan protocol (633nm excitation, neutral density filter). Following salt scan, free *N*-hydroxysuccinimide activated groups were blocked with 50mM ethanolamine in 50mM borate buffer at a final pH of 9.2. Slides were then rinsed with deionized water, dried and stored at room temperature in a desiccator.

Sample preparation. Each serum sample was diluted 1:15 with PBS containing 0.1% v/v Tween20 and 3% w/v BSA, thoroughly vortexed for 15 seconds and incubated at 37°C for 15 minutes to dissolve potential lipid aggregates. Insoluble residual sample components were centrifuged for 30 seconds in a table-top centrifuge at maximum speed. Samples were transferred to the array slides and gently rocked in a sealed humidified incubator for 2h at 37°C. Unbound sample components were washed with a series of 0.1% and 0.001% Tween 20 in PBS. ImmunoPure goat anti-human IgA + IgG + IgM conjugated to long chain biotin ("Combo", Pierce, Rockford, IL) 1:100 in PBS containing 0.1% Tween20 and 3% BSA was added and slides were incubated at room temperature in a humidified chamber for 45 minutes. Following washing steps as described above, bound antibodies were visualized by incubating slides with fluorescent dye streptavidin Alexa Fluor^{555} (Molecular Probes, Invitrogen, Carlsbad, CA), 1:1000 in PBS/0.1% Tween 20 at room temperature for 30 minutes.

Data quantification. Fluorescence signals corresponding to glycan-bound antibodies were measured and quantification of images was carried out using ImaGene analysis software version 6.1 and 7.5 (BioDiscovery, El Segundo, US). Salt scan and fluorescence image of each slide were aligned to assure quantification accuracy in case of weak signals. Signals were measured as medTSI per glycan and were expressed as median across eight within-array replicates.

Statistical analysis. Data analysis was performed using the open source statistical programming language R (http://CRAN.R-project.org/, version 2.8.1). All p-values < 0.05 were taken as significant. Epitopic and structural preference of antibody binding to individual carbohydrate structures was carried out using hierarchical clustering. We performed

unsupervised agglomerative hierarchical clustering using Ward's method[31] with 1-correlation distance on the signals from the 50µM prints, including both the 203 glycans and the 8 biotin controls. Cancer-specific glycans were identified by univariate linear modeling implemented in the R package limma[32]. For multivariate classification, sets of glycans were selected by a support vector machine algorithm (R package GALGO)[33] ranked by their stability within 700 independent statistical runs. The most stable glycans were further combined in a linear discriminant model (R package 'MASS')[34] and analyzed by ROC curves (R package ROCR)[35] for each binary classifier to determine the sensitivity and specificity of selected models. These models were further used to compare glycan-based classifiers to the standard tumor marker CA125. The best cut-off between observed false negative and false positive rates was determined with two independent methods: "best cut-off point" (Method A), defined as the largest value of sensitivity plus specificity, selected from each individual classifier in ROC curves; and "precision-recall break even point" (Method B)[35], the point at which precision equals recall and predictions are made due to the prevalence within given data.

Variability and reproducibility. Quality control intra-chip analysis using replicates and inter-chip reproducibility were previously determined as described[20]. To assess the experimental variability of the array, samples from 32 patients (13 controls and 19 tumors) were randomly chosen for replicate profiling in two independent experiments on different days. In order to detect technology based systematic errors, mean, s.d., and CV were assessed.

Biological and technical controls. As biological controls, α-rhamnose and aminoglucitol were used. Alpha-rhamnose was used as a "positive biological control" due to the known high expression levels of anti-rhamnose antibodies in healthy control individuals[21]. Aminoglucitol, an opened reduced form of *D*-glucose which is not present as a structural component of regular glycosylation, has been shown to be negative for anti-glycan binding at similar values to the technical background binding control[20], which is why we used it as a negative biological control.

3.2.4 Results

Reproducibility of the printed glycan array. Comparing independent experimental settings, we found a high concordance correlation coefficient (CCC) for the mean (across patients) of the median total signal intensity (medTSI) of eight within-chip replicates (CCC=0.957; Figure 12 A). Standard deviations (CCC=0.940; Figure 12 B) within all measured samples with the coefficient of variation (CV) were concordant to these findings (CCC=0.966; Figure 12 C). The CV across mean signal intensities for repeats had a mean of 14.6%, with a

median of 14.2% and interquartile range 8.9% - 19.0% (Figure 12 D). Two independent experimental repeats showed a similar maximal CV for the same oligosaccharides: *NeuAcα2-3Galβ1-3(NeuAcα2-6)GalNAcα* ($CV_{exp.1}$ 433.3%; $CV_{exp.2}$ 463.6%), *Galβ1-4GlcNAcβ1-3(Galβ1-4GlcNAcβ1-6)Galβ1-4GlcNAc* ($CV_{exp.1}$ 442.8%; $CV_{exp.2}$ 371.9%) and $(SA2\text{-}6G\text{-}GN\text{-}M)_2$-3,6-M-GN-GNβ ($CV_{exp.1}$ 415.2%; $CV_{exp.2}$ 393.8%; Figure 12 C). The "biological positive control" α-rhamnose showed a $CV_{exp.1}$ 49.7% and $CV_{exp.2}$ 46.8%, respectively. We examined intra-slide variability using the carbohydrate structure *Fucα1-2(GalNAcβ1-3)Galβ*, which was represented twice on the array in two different positions and had a CV of less than 10% in each experiment ($CV_{exp.1}$ 8.7%; $CV_{exp.2}$ 9.9%).

AB0 blood group antigens as proof of principle. Anti-A and anti-B blood group antibody binding to corresponding AB0-blood group antigens printed onto the array was used as an attempted validation. Investigated ligands for blood group antibodies were A and B di-, tri- and tetrasaccharides. Corresponding antibody patterns in sera of patients with known blood groups were defined by standardized clinical agglutination tests and compared to array-based results. The highest levels of statistical significance ($p<0.05$ for 19 glycans) were detected for A and B tri- and tetrasaccharides. Two printed trisaccharide structures known to be minimal blood group determinants, *GalNAcα1-3(Fucα1-2)Galβ,* A trisaccharide/ A_{tri} and *Galα1-3(Fucα1-2)Galβ,* B trisaccharide/ B_{tri} showed a highly significant difference ($p<0.001$). Low or no signals were detected in the case of A_{tri} in blood groups A and AB and against B_{tri} in blood groups B and AB. In contrast, high levels of antibodies were found for both trisaccharides in blood group O, for A_{tri} in blood group B and for B_{tri} in blood group A (Figure 13). The most significant difference in antibody levels ($p<0.001$) was observed for the tetrasaccharides *GalNAcα1-3(Fucα1-2)Galβ1-4GlcNAcβ* ($A_{type\ 2}$) and *Galα1-3(Fucα1-2)Galβ1-4GlcNAcβ* ($B_{type\ 2}$), known to be major AB0 antigens on erythrocytes. In addition to expected differences in antibody profiles to $A_{type\ 2}$ in blood groups O and B *versus* A and AB, it was noted that the level of antibodies differed highly significantly between blood groups B and O ($p<0.001$). A similar result was found for $B_{type\ 2}$ with low antibody levels in blood groups AB and B compared to A and O, which significantly differed from lower antibody levels in blood groups A *versus* O ($p=0.005$). Differentiating antibody binding across blood groups was not possible using the disaccharides *GalNAcα1-3Galβ* (A_{di}; $p= 0.0841$) and *Galα1-3Galβ* (B_{di}; $p= 0.9189$), which are lacking fucose residues. Although high levels of antibody binding could be detected to A_{di} with lower signals against B_{di} across all profiled blood groups (Figure 13), the highest levels were found in blood group O individuals against A and B tri- and tetrasaccharides.

Anti-glycan antibody distribution across printed glycans. The highest anti-glycan antibody levels (mean ± s.d. in medTSI) were observed for glycan *3'-O-Su-Galβ1-3GlcNAcβ* ($146*10^5 \pm 68*10^5$), GlcNAcβ1-*4Mur-L-Ala-D-i-Gln-Lys* ($94.6*10^5 \pm 53*10^5$) and *Neu5Acα2-3Galβ1-3GlcNAcβ* ($87*10^5 \pm 37*10^5$) (Figure 14). The "biological positive control" α-rhamnose ($50*10^5 \pm 24*10^5$; Figure 14 A) showed high signals whilst the "negative control" aminoglucitol ($2.5*10^5 \pm 4.2*10^5$; Figure 14 C) had a low mean result. The glycan which was examined twice on the array to analyze any intra-slide variability (*Fucα1-2(GalNAcβ1-3)Galβ*; Figure 14 B) showed close levels in both analyses.

A large spectrum of anti-glycan antibody signals was detected within the whole patient cohort, composing of healthy controls, non-mucinous borderline tumors and ovarian cancers (Table 6). The mean medTSI for all carbohydrate-bound antibody levels per patient varied between $8.59*10^5$ and $38.21*10^5$ medTSI. The interquartile range varied from $8.08*10^5$ to $49.7*10^5$ medTSI and the CV from 79.7% to 188.4%. All individual serum samples showed mild and extreme outliers (Figure 15). Observed biological variability of anti-glycan antibodies was hereby not associated with any clinical or experimental covariates, namely antibody levels in cancer patients were not as a whole altered as compared to healthy control patients.

Anti-glycan antibodies bind to specific epitope core structures. Cluster analysis revealed 53 clusters, each containing at most eight oligosaccharides that show similarities in carbohydrate structures. Glycans that are structurally identical but printed using different spacers, clustered together, as did the eight biotin controls and the blood group A and B antigens. Distinct structural similarities were found for 21 clusters (Table 7): *N*-linked glycans (11 clusters), *O*-linked glycans (3 clusters), glycosphingolipids (2 clusters) and not clearly sub-specifiable structures (5 clusters). <u>*N*-linked carbohydrates</u> included a. linear lactosamine structures [*Galβ1-4GlcNAcβ*]$_{1-3}$; b. linear glucosamine structures [*GlcNAcβ1-4*]$_{3-6}$; c. sulfated lactosamine structures *4'* or *6'-O-Su Galβ1-4GlcNAcβ*; d. single core lactosamine structures *Galβ1-4GlcNAcβ* with terminally coupled "monsters" [20], representing artificial carbohydrates which are not present in biological objects, namely *β2-6Neu5Gc*, *β2-6Neu5Ac* or *Galα1-4Galβ1-4GlcNAcβ*; e. single core lactosamine structures *Galβ1-4GlcNAcβ* modified by *Galα* or O-sulfation; f. Lewis c (*Galβ1-3GlcNAcβ*) structures; g. single core lactosamine structures modified by neuraminic acid (*NeuAcα2-3Galβ1-4GlcNAcβ*; h. single core lactosamine structures with additional *N*-acetylglucosamine (*GlcNAcβ1-6Galβ1-4GlcNAcβ*; and i. single core *N*-actelylglucosamine structures (*GlcNAcβ*). Interestingly, within this cluster four out of eight carbohydrate structures (*Fucβ1-3GlcNAcβ, Galβ1-3GlcNAcβ*, *3'-O-Su-Galβ1-3GlcNAcβ* and *NeuAcα2-6Galβ1-3GlcNAcβ*) were highly correlated in their signal intensities, with

correlation coefficients ranging from 0.76 to 0.90. *O*-linked carbohydrates included: a. Thomson-Friedensreich structures *Galβ1-3GalNAcα*; b. sialylated Tn antigen structures *Neu5Acα2-6GalNAcα*; c. lactosamine on Tn antigen *Galβ1-4GlcNAcβ1-3GalNAcα*. The last specific structure group consisted of glycosphingolipids including *Galβ1-3GlcNAcβ1-3Galβ1-3Glcβ*.

Non-mucinous ovarian cancer biomarker detection. Linear modeling revealed 24 carbohydrate structures for which the amount of anti-glycan antibodies was significantly lower in the non-mucinous ovarian borderline and cancer cohort as compared to the healthy patient cohort. The glycan structure with the most significant discriminatory ability was P_1 (*Galα1-4Galβ1-4GlcNAcβ*; p=0.0008). Compared to the presently used tumor marker CA125, which had a sensitivity of 64/76% (Method A/B) and specificity of 95.6/73.9% (Method A/B) in our cohort, the identified anti-glycan antibodies revealed comparable values, particularly the top candidate P_1 (sensitivity 79.2/70.8% (Method A/B) and specificity 75.7/78.8% (Method A/B)) (Figure 16). A combination of both P_1 and CA125 did improve neither sensitivity nor specificity (82.6/76.0% (Method A/B) and 72.0/73.9% (Method A/B), respectively).

We next aimed to identify a panel of anti-glycan antibodies that could generate a higher sensitivity and specificity for the differentiation between non-mucinous ovarian borderline and cancer patients from healthy control patients. Interestingly, the top candidate P_1 was again the most stable glycan using this selection algorithm. In comparison to the univariate approach, the first three multivariate selected and linear combined glycans (svmLDA03) improved the sensitivity by 5-10% and the specificity by 0-3% (Table 8). Combination of a panel of six glycans generated the highest sensitivity (79.2%; Method A/B) and specificity (84.8%; Method A/B): *Galα1-4Galβ1-4GlcNAcβ*, *Neu5Acα2-3Galβ1-4-(6-Su)GlcNAcβ*, *Neu5Acβ2-6GalNAcα*, *Neu5Acβ2-6Galβ1-4GlcNAcβ*, *Neu5Acα2-3Galβ1-3-(6-Su)GalNAcα* and *Galβ1-4GlcNAcβ1-6Galβ1-4GlcNAcβ* (svmLDA06) (Figure 16; Table 8). The combination of the highest significant ten glycans in the multivariate selection model (svmLDA10) only improved sensitivity by 4% (Method A) (Table 8). Interestingly, these ten glycans share similar carbohydrate motifs, namely eight structures containing core single lactosamine (*Galβ1-4GlcNAcβ*) and two structures with Tn antigen (*GalNAcα*) (Table 9). In contrast, different monosaccharides, *e.g. Galα-*, *Fucα-* as well as sulfated and sialylated structures were observed at the terminal regions of these ten glycans. So called "Monster" structures presented as *Neu5Acβ2-6* were also observed, but do not seem to play a biological role compared to core structures.

3.2.5 Discussion

Using printed glycan array technology, we identified patterns of core carbohydrate structures in ovarian cancer patients which are specifically recognized by anti-glycan antibodies. Most glycan binding proteins are able to recognize glycan epitopes[36], and non-glycan moiety of the glycoconjugate is discussed to be involved in this binding. This is concordant with our findings, as we identified certain core structures which seem to recognize specific anti-glycan antibodies. Based on this observation we assume that the core of identified carbohydrate structures, as well as terminal or peripheral sugars, could play an important role in antibody recognition. The binding affinity or specificity of anti-glycan antibodies seems to be higher in the case of specific core structures[20], especially containing *GalNAcα-*, *Galβ1-4GlcNAc-* or *Galβ1-4Glcβ-*. Potentially, benign human cell core structures could be occupied by interactions of different types of membrane- associated proteins as well as other carbohydrate structures. Therefore, binding of anti-glycan antibodies seem to be limited to these potential cell surface carbohydrate epitopes. During oncogenic transformation of cells and rearrangements of extra-cellular matrix, these encrypted binding sites could become available for circulating anti-glycan antibodies. This means that glycans and their molecular environment are presenting epitopes and signal the status of a cell to the immune system leading to a corresponding response. Various functional studies have been performed in the field of detection of naturally occurring and anti-glycan antibodies and their carbohydrate-binding, *e.g.* for anti-α-galactosyl antibodies[37,38], antibodies bound to negatively charged glycans[20,39], or *GlcNAc*-terminated carbohydrate chains[20,40]. This could reflect a natural antigenic mimicry, described as anti-idiotypic antibodies[41] which are peptides that mimic tumor-associated carbohydrate antigens or their glycosphingolipid constituent.

At present no technologies are available to study anti-glycan antibody reactions as sensitive and reproducible as within this high-throughput glycan array technology. Other technologies as ELISA and suspension array are under evaluation, but the main advantage of an array is obviously the unlimited number of glycans which can be screened in a high-throughput manner. We have therefore used the well established AB0 blood group carbohydrate antigens as an attempted method to validate the glycan array results. AB0-specific antibodies (isohemagglutinins) belong to a group of naturally occurring antibodies which can be determined by hemagglutination, ELISA[42] and FACS[43]. Although in our experimental setting anti-glycan antibodies were determined using a pool of immunoglobulin subtypes, our results on blood groups detected by glycan array are concordant with classical hemagglutination data. However, only anti-tri/tetra-saccharides, and no disaccharides, were

able to discriminate within blood groups, whilst anti-disaccharide antibodies seemed to be less specific and bound to corresponding blood group antigens independent of AB0 blood groups, as previously described[20,24]. In contrast, A/B tri- and tetrasaccharides showed highly specific binding with appropriate antibodies in concordance with basic distribution of anti- A/B antibodies within blood groups. Affinity of AB0 blood group antibodies increased due to the number of sugar units in A/B antigens, being lowest in A/B disaccharides and highest in A/B tetrasaccharides, with fewer outliers observable in longer carbohydrate structures. Some degree of auto-reactivity could be found in our experiments in concordance with the literature, probably due to the heterogeneity of the A and B blood groups[43]. These findings on the AB0 blood group system, although used as an attempted validation is the first published experiments on the AB0 blood group system using a glycan array approach.

Since there are no carbohydrates that are yet established as truly positive and negative antibody binding reference structures, α-rhamnose is used in this study as positive and Aminoglucitol as a negative biological control. The overall variability of signals for anti-glycan antibodies bound to carbohydrate antigens measured in this study are similar to the literature as is the coefficient of variation for the proposed positive biological control *α-rhamnose*[24,44]. Minor discrepancies can be explained by different technical set-ups, including a. differences in carbohydrate presentation on individual platforms, as monomeric structures on bovine serum albumin coupled to a platform[24,27] *versus* monomeric structures coated covalently to a surface matrix[20,28] ; b. differences in generation of individual serum samples; and c. one-step detection[24,27,28] *versus* two-step detection methodology[20]. Although the amount and concentration of spotted glycans can also influence the data, the printed glycan array has shown high reproducibility in measuring anti-glycan antibodies bound to spotted carbohydrate structures[20,24].

Ovarian cancer biomarker discovery using glycan array. Previously, the printed glycan array was used to study immunoprofiles of serum anti-glycan antibodies in 106 healthy female blood donors[20], detecting a broad variety of antibody profiles, independent of clinical parameters. Although similar individual anti-glycan antibody patterns have been shown in another group of patients[27], other studies have observed discrepancies due to differing age groups[45] or varying antibody titres against the carbohydrate *mannan* of *Candida albicans,* causing variations as much as 6.000-fold within a healthy donor group[46,47]. Therefore, interpretations on one single cohort which moreover has no information regarding their clinico-pathological characteristics, could harvest mistakes. In this study we have therefore used the printed glycan array as a method to identify particular anti-glycan antibodies discriminating

between the clinically most important screening setting, namely non-mucinous ovarian borderline and cancer patients from unaffected women. P_1, a member of the P blood group system demonstrated the highest significance and similar sensitivity/specificity levels as CA125 in detecting non-mucinous ovarian tumors in this setting. The P blood group system has two common phenotypes, P_1 and P_2, P_1 and P^k (CD77; *Galα1-4Galβ1-4Glc*) share a common structure, namely *Galα1-4Galβ1-4Glc(NAc)*. The physiological function of P_1 and the P blood group system in general is still unknown. P^k antigen has been associated with acute leukaemia[48], whilst P_1 is present in hematopoietic[49] and mesothelioma cell lines[50]. The presence of P_1 in epithelial cells has been correlated to urinary tract infections due to the terminal part of P_1, *Galα1-4Gal,* being a bacterial adhesion molecule in a pyelonephrotic strain of *Escherichia coli*[51]. The detection of P_1 within this experimental setting as a tumor-associated carbohydrate antigen in ovarian cancers could implement that similarly to mesothelioma cells also ovarian cancer cells contain a P_1 epitope. In this case the lowered levels of anti-P_1 antibodies in the serum of cancer patients could be explained due to the higher amount of P_1 cell-bound antibodies.

Except for P_1, no other anti-glycan antibody in this experimental approach has reached a similar performance as CA125. This could be due to the current limited panel of synthetic carbohydrates within the printed glycan array. It certainly reflects the heterogeneity even in a non-mucinous ovarian cancer cohort, and further clear-cut cohort selections will be necessary to improve sensitivities and specificities. Further, a panel of markers as suggested in proteogenomic studies[52,53] and as confirmed in our data with already the combination of two anti-glycan antibodies improved sensitivity and specificity compared to the identified univariate candidates, including the top candidate P_1. An improved differentiation could be reached with a minimum of six selected glycans with only minor gains by combining an even higher number of glycans. Panels of anti-glycan antibodies did not, however, reach a better sensitivity and specificity than 84%. When using a combination of eight candidates containing the structure *Galβ1-4GlcNAcβ* in the top ten selected antibodies, a sensitivity of 83.3% and specificity of 87.8% could be reached, which is clearly better than CA125 (76%/ 73.9%, respectively). Possible explanations why sensitivities and specificities above 90% could not be reached are a. the heterogeneity of the patient cohort; b. unknown clinically relevant diseases in the healthy controls; c. background binding effects which are not necessarily cancer-associated; d. low-titer reactivities of potential tumor-associated signals; e. vast variety of serum anti-glycan antibodies expressing cross-reactivity, and f. limited number of synthesized glycans on the array. We also need to be aware that using anti-glycan antibody detection *via* array

technology cannot offer information about inducing stimuli, neither are we able to define the affinity or specificity of a collective of antibodies binding to individual glycan structures. A positive binding signal of human immunoglobulins to glycans demonstrates therefore only a variety of polyclonal antibodies binding to a saccharide that possibly represents different forms of biological epitopes.

3.2.6 Conclusions

Aberrations in well programmed glycosylation occur very early during malignant transformation and result in the appearance of specific glycans known as tumor-associated carbohydrate antigens on cell surface and serum components. They are also known to trigger an immune response resulting in the generation of anti-glycan antibodies. This is the first study publishing the identification of patient-specific anti-glycan antibody patterns in non-mucinous ovarian cancer. We found a discriminating anti-glycan antibody panel including the top candidate P_1 which can diagnose ovarian borderline tumors and cancers with higher sensitivity and specificity than CA125. Except for P_1, no other individual anti-glycan antibody in this experimental approach has reached a similar performance as CA125. This could be due to the current limited panel of synthetic carbohydrates within the printed glycan array. It certainly also reflects the heterogeneity even in a non-mucinous ovarian cancer cohort, and further clear-cut cohort selections will be necessary to improve sensitivity and specificity studies. Printed glycan array technology has the potential to detect individual and panels of anti-glycan antibodies and is therefore a highly promising tool for future ovarian cancer biomarker discovery.

3.2.7 Acknowledgements

This work was supported by the Cancer League of the Canton of Zurich (V.H.S.); Oncosuisse (02115-08-2007 to V.H.S.); Swiss National Foundation (320000-12543 to V.H.S.), European Science Foundation (V.H.S.); Cellexicon, Inc. La Jolla, CA, USA; National Cancer Institute, USA, grant 1U01CA128526 (to M.E.H).

3.2.8 References

1. Ozols, R. F., Challenges for chemotherapy in ovarian cancer. *Ann Oncol* **2006,** 17 Suppl 5, v181-7.
2. Fishman, D. A.; Bozorgi, K., The scientific basis of early detection of epithelial ovarian cancer: the National Ovarian Cancer Early Detection Program (NOCEDP). *Cancer Treat Res* **2002,** 107, 3-28.

3. Sasaroli, D.; Coukos, G.; Scholler, N., Beyond CA125: the coming of age of ovarian cancer biomarkers. Are we there yet? *Biomark Med* **2009,** 3, (3), 275-288.
4. Jacob, F.; Goldstein, D. R.; Fink, D.; Heinzelmann-Schwarz, V., Proteogenomic studies in epithelial ovarian cancer: established knowledge and future needs. *Biomarkers Med.* **2009,** 3, (6), 743-756.
5. Hakomori, S., Tumor-associated carbohydrate antigens defining tumor malignancy: basis for development of anti-cancer vaccines. *Adv Exp Med Biol* **2001,** 491, 369-402.
6. Hakomori, S., Glycosylation defining cancer malignancy: new wine in an old bottle. *PNAS* **2002,** 99, (16), 10231-3.
7. Gyorgy, B.; Tothfalusi, L.; Nagy, G.; Pasztoi, M.; Geher, P.; Lorinc, Z.; Polgar, A.; Rojkovich, B.; Ujfalussy, I.; Poor, G.; Pocza, P.; Wiener, Z.; Misjak, P.; Koncz, A.; Falus, A.; Buzas, E. I., Natural autoantibodies reactive with glycosaminoglycans in rheumatoid arthritis. *Arthritis Res Ther* **2008,** 10, (5), R110.
8. Dotan, I.; Fishman, S.; Dgani, Y.; Schwartz, M.; Karban, A.; Lerner, A.; Weishauss, O.; Spector, L.; Shtevi, A.; Altstock, R. T.; Dotan, N.; Halpern, Z., Antibodies against laminaribioside and chitobioside are novel serologic markers in Crohn's disease. *Gastroenterology* **2006,** 131, (2), 366-78.
9. Fuster, M. M.; Esko, J. D., The sweet and sour of cancer: glycans as novel therapeutic targets. *Nat Rev Cancer* **2005,** 5, (7), 526-42.
10. Lutz, H. U.; Binder, C. J.; Kaveri, S., Naturally occurring auto-antibodies in homeostasis and disease. *Trends Immunol* **2009,** 30, (1), 43-51.
11. Schwartz-Albiez, R.; Laban, S.; Eichmuller, S.; Kirschfink, M., Cytotoxic natural antibodies against human tumours: an option for anti-cancer immunotherapy? *Autoimmun Rev* **2008,** 7, (6), 491-5.
12. Huang, C. Y.; Thayer, D. A.; Chang, A. Y.; Best, M. D.; Hoffmann, J.; Head, S.; Wong, C. H., Carbohydrate microarray for profiling the antibodies interacting with Globo H tumor antigen. *Proc Natl Acad Sci U S A* **2006,** 103, (1), 15-20.
13. Huflejt, M. E.; Blixt, O.; Vuskovic, M.; Xu, H.; Shaw, L. E.; Reuben, J. M.; Kuerer, H. M.; Cristofanilli, M. In *Glycan array identifies specific signatures of anti-glycan autoantibodies in sera of breast cancer patients: diagnostic, prognostic and therapeutic opportunities.*, 28th Annual San Antonia Breast Cancer Symposium, San Antonia, TX, 2005; Breast Cancer Res. Treat.: San Antonia, TX, 2005; p 85.
14. Lawrie, C. H.; Marafioti, T.; Hatton, C. S.; Dirnhofer, S.; Roncador, G.; Went, P.; Tzankov, A.; Pileri, S. A.; Pulford, K.; Banham, A. H., Cancer-associated carbohydrate identification in Hodgkin's lymphoma by carbohydrate array profiling. *Int J Cancer* **2006,** 118, (12), 3161-6.
15. Shin, I.; Park, S.; Lee, M. R., Carbohydrate microarrays: an advanced technology for functional studies of glycans. *Chemistry* **2005,** 11, (10), 2894-901.
16. Manimala, J. C.; Roach, T. A.; Li, Z.; Gildersleeve, J. C., High-throughput carbohydrate microarray analysis of 24 lectins. *Angew Chem Int Ed Engl* **2006,** 45, (22), 3607-10.
17. Moller, I.; Marcus, S. E.; Haeger, A.; Verhertbruggen, Y.; Verhoef, R.; Schols, H.; Ulvskov, P.; Mikkelsen, J. D.; Knox, J. P.; Willats, W., High-throughput screening of monoclonal antibodies against plant cell wall glycans by hierarchical clustering of their carbohydrate microarray binding profiles. *Glycoconj J* **2008,** 25, (1), 37-48.
18. Disney, M. D.; Seeberger, P. H., The use of carbohydrate microarrays to study carbohydrate-cell interactions and to detect pathogens. *Chem Biol* **2004,** 11, (12), 1701-7.
19. Park, S.; Shin, I., Carbohydrate microarrays for assaying galactosyltransferase activity. *Org Lett* **2007,** 9, (9), 1675-8.

20. Huflejt, M. E.; Vuskovic, M.; Vasiliu, D.; Xu, H.; Obukhova, P.; Shilova, N.; Tuzikov, A.; Galanina, O.; Arun, B.; Lu, K.; Bovin, N., Anti-carbohydrate antibodies of normal sera: findings, surprises and challenges. *Mol Immunol* **2009,** 46, (15), 3037-49.
21. Dotan, N.; Altstock, R. T.; Schwarz, M.; Dukler, A., Anti-glycan antibodies as biomarkers for diagnosis and prognosis. *Lupus* **2006,** 15, (7), 442-50.
22. Krishnamoorthy, L.; Mahal, L. K., Glycomic analysis: an array of technologies. *ACS Chem Biol* **2009,** 4, (9), 715-32.
23. Bovin, N. V.; Huflejt, M. E., Unlimited glycochip. *Trends Glycosci. Glycotechnol.* **2008,** 20, (115), 245-258.
24. Oyelaran, O.; McShane, L. M.; Dodd, L.; Gildersleeve, J. C., Profiling Human Serum Antibodies with a Carbohydrate Antigen Microarray. *J Proteome Res* **2009**.
25. Blixt, O.; Hoffmann, J.; Svenson, S.; Norberg, T., Pathogen specific carbohydrate antigen microarrays: a chip for detection of Salmonella O-antigen specific antibodies. *Glycoconj J* **2008,** 25, (1), 27-36.
26. Wang, D.; Carroll, G. T.; Turro, N. J.; Koberstein, J. T.; Kovac, P.; Saksena, R.; Adamo, R.; Herzenberg, L. A.; Herzenberg, L. A.; Steinman, L., Photogenerated glycan arrays identify immunogenic sugar moieties of Bacillus anthracis exosporium. *Proteomics* **2007,** 7, (2), 180-4.
27. Oyelaran, O.; Li, Q.; Farnsworth, D.; Gildersleeve, J. C., Microarrays with varying carbohydrate density reveal distinct subpopulations of serum antibodies. *J Proteome Res* **2009,** 8, (7), 3529-38.
28. Wang, C. C.; Huang, Y. L.; Ren, C. T.; Lin, C. W.; Hung, J. T.; Yu, J. C.; Yu, A. L.; Wu, C. Y.; Wong, C. H., Glycan microarray of Globo H and related structures for quantitative analysis of breast cancer. *Proc Natl Acad Sci U S A* **2008,** 105, (33), 11661-6.
29. Turnbull, J. E.; Field, R. A., Emerging glycomics technologies. *Nat Chem Biol* **2007,** 3, (2), 74-7.
30. Blixt, O.; Head, S.; Mondala, T.; Scanlan, C.; Huflejt, M. E.; Alvarez, R.; Bryan, M. C.; Fazio, F.; Calarese, D.; Stevens, J.; Razi, N.; Stevens, D. J.; Skehel, J. J.; van Die, I.; Burton, D. R.; Wilson, I. A.; Cummings, R.; Bovin, N.; Wong, C. H.; Paulson, J. C., Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. *Proc Natl Acad Sci U S A* **2004,** 101, (49), 17033-8.
31. Ward, J. H., Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association* **1963,** 58, 236-244.
32. Smyth, G. K., Statistical Applications in Genetics and Molecular Biology. In *Linear models and empirical Bayes methods for assesing differential expression in microarray experiments.*, 3, Ed. 2004; pp 68, 232, 397, 399, 416, 418, 436.
33. Trevino, V.; Falciani, F., GALGO: an R package for multivariate variable selection using genetic algorithms. *Bioinformatics* **2006,** 22, (9), 1154-6.
34. Venables, W. N.; Ripley, B. D., *Modern Applied statistics with S*. Springer, New York: New York, 2002.
35. Sing, T.; Sander, O.; Beerenwinkel, N.; Lengauer, T., ROCR: visualizing classifier performance in R. *Bioinformatics* **2005,** 21, (20), 3940-1.
36. Varki, A.; Cummings, R.; Esko, J.; Freeze, H.; Hart, G.; Marth, J., *Essentials of Glycobiology*. Cold Spring Harbor Labartory Press: Plainview (NY), 1999; p p.41.
37. Milland, J.; Yuriev, E.; Xing, P. X.; McKenzie, I. F.; Ramsland, P. A.; Sandrin, M. S., Carbohydrate residues downstream of the terminal Galalpha(1,3)Gal epitope modulate the specificity of xenoreactive antibodies. *Immunol Cell Biol* **2007,** 85, (8), 623-32.
38. Parker, W.; Lin, S. S.; Yu, P. B.; Sood, A.; Nakamura, Y. C.; Song, A.; Everett, M. L.; Platt, J. L., Naturally occurring anti-alpha-galactosyl antibodies: relationship to xenoreactive anti-alpha-galactosyl antibodies. *Glycobiology* **1999,** 9, (9), 865-73.

39. Lekakh, I. V.; Bovin, N. V.; Bezyaeva, G. P.; Poverenny, A. M., Natural hidden autoantibodies react with negatively charged carbohydrates and xenoantigen Bdi. *Biochemistry (Mosc)* **2001,** 66, (2), 163-7.
40. Briko, N. I.; Bovin, N. V.; Shevelev, B. I.; Dynga, L. O.; Blinnikova, Y. I.; Kuksyuk, P. P.; Myasoyedova, S. E.; Ambrosov, I. V.; Filatov, N. N., Enzyme immunoassay kit for detecting antibodies to group-specific antigen of group A Streptococcus on the base of conjugated N-acetylglucosamine and its medical application. *Klin.La.Diagn.* **1997,** 9, 43-46.
41. Chapman, P. B., Anti-idiotypic monoclonal antibody cancer vaccines. *Semin Cancer Biol* **1995,** 6, (6), 367-74.
42. Buchs, J. P.; Nydegger, U. E., Development of an ABO-ELISA for the quantitation of human blood group anti-A and anti-B IgM and IgG antibodies. *J Immunol Methods* **1989,** 118, (1), 37-46.
43. Stussi, G.; Huggel, K.; Lutz, H. U.; Schanz, U.; Rieben, R.; Seebach, J. D., Isotype-specific detection of ABO blood group antibodies using a novel flow cytometric method. *Br J Haematol* **2005,** 130, (6), 954-63.
44. Schwarz, M.; Spector, L.; Gortler, M.; Weisshaus, O.; Glass-Marmor, L.; Karni, A.; Dotan, N.; Miller, A., Serum anti-Glc(alpha1,4)Glc(alpha) antibodies as a biomarker for relapsing-remitting multiple sclerosis. *J Neurol Sci* **2006,** 244, (1-2), 59-68.
45. Faux, J. A.; Agbarakwe, A. E.; Misbah, S. A.; Chapel, H. M., A comparison of specific IgG antibody levels to the cell wall mannan of Candida albicans in normal individuals and in patients with primary antibody deficiency. *J Immunol Methods* **1992,** 153, (1-2), 167-72.
46. Lehmann, P. F.; Reiss, E., Comparison by ELISA of serum anti-Candida albicans mannan IgG levels of a normal population and in diseased patients. *Mycopathologia* **1980,** 70, (2), 89-93.
47. Greenfield, R. A.; Stephens, J. L.; Bussey, M. J.; Jones, J. M., Quantitation of antibody to Candida mannan by enzyme-linked immunosorbent assay. *J Lab Clin Med* **1983,** 101, (5), 758-71.
48. Brodin, N. T.; Dahmen, J.; Nilsson, B.; Messeter, L.; Martensson, S.; Heldrup, J.; Sjogren, H. O.; Lundblad, A., Monoclonal antibodies produced by immunization with neoglycoproteins containing Gal alpha 1-4Gal beta 1-4Glc beta-O and Gal alpha 1-4Gal beta 1-4GlcNAc beta-O residues: useful immunochemical and cytochemical reagents for blood group P antigens and a differentiation marker in Burkitt lymphoma and other B-cell malignancies. *Int J Cancer* **1988,** 42, (2), 185-94.
49. Dunstan, R. A.; Simpson, M. B.; Rosse, W. F., Presence of P blood group antigens on human platelets. *Am J Clin Pathol* **1985,** 83, (6), 731-5.
50. Spitalnik, S. L.; Spitalnik, P. F.; Dubois, C.; Mulshine, J.; Magnani, J. L.; Cuttitta, F.; Civin, C. I.; Minna, J. D.; Ginsburg, V., Glycolipid antigen expression in human lung cancer. *Cancer Res* **1986,** 46, (9), 4751-5.
51. Lomberg, H.; Jodal, U.; Eden, C. S.; Leffler, H.; Samuelsson, B., P1 blood group and urinary tract infection. *Lancet* **1981,** 1, (8219), 551-2.
52. Moore, R. G.; Brown, A. K.; Miller, M. C.; Skates, S.; Allard, W. J.; Verch, T.; Steinhoff, M.; Messerlian, G.; DiSilvestro, P.; Granai, C. O.; Bast, R. C., Jr., The use of multiple novel tumor biomarkers for the detection of ovarian carcinoma in patients with a pelvic mass. *Gynecol Oncol* **2008,** 108, (2), 402-8.
53. Mor, G.; Visintin, I.; Lai, Y.; Zhao, H.; Schwartz, P.; Rutherford, T.; Yue, L.; Bray-Ward, P.; Ward, D. C., Serum protein markers for early detection of ovarian cancer. *Proc Natl Acad Sci U S A* **2005,** 102, (21), 7677-82.

3.2.9 Figures

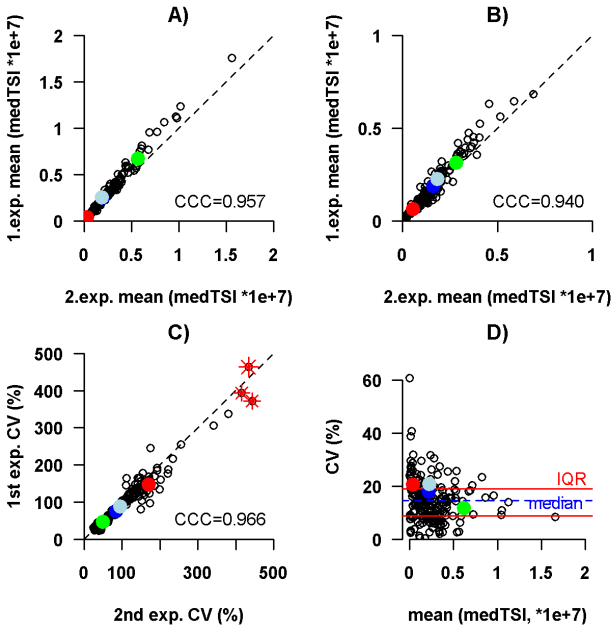


Figure 12 Overall inter-slide variability.
Results of two independent printed glycan array experiments on 32 patients for A. means; B. standard deviation (s.d.); C. coefficient of variation; D. summarized coefficient of variation. Number in each plot refers to concordance correlation coefficient (CCC), red line indicates IQR, blue dashed line indicates median, green dot indicates α-rhamnose; red dot indicates aminoglucitol, light and dark blue dot indicate *Fucα1-2(GalNAcβ1-3)Galβ* printed in double on different positions across the array. *NeuAcα2-3Galβ1-3(NeuAcα2-6)GalNAcα*, *Galβ1-4GlcNAcβ1-3(Galβ1-4GlcNAcβ1-6)Galβ1-4GlcNAc* and *(SA2-6G-GN-M)$_2$-3,6-M-GN-GNβ* are highlighted as carbohydrate structures with highest coefficient of variation (red star; C).

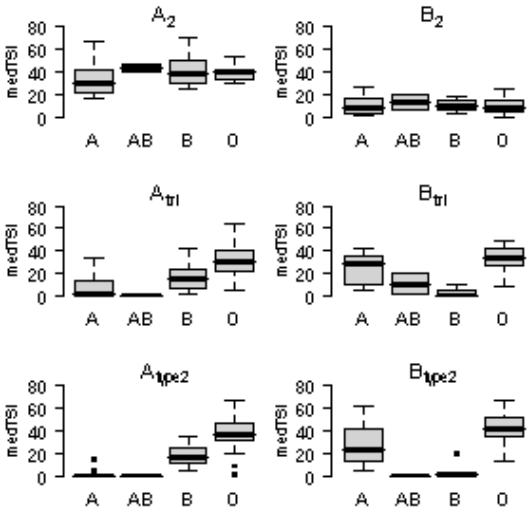


Figure 13 Validation using AB0 blood group anti-glycan antibody detection.
Boxplots demonstrate the distribution of medTSI within each blood group independent of disease status, demonstrating the binding of anti-glycan antibodies against A and B blood group antigens. Comparisons were measured using the Kruskal-Wallis test. Carbohydrate structures presented are: A_{di} p=0.064; B_{di} p=0.898; A_{tri} p<0.0001; B_{tri} p<0.0001; A_{tetra} p<0.0001; B_{tetra} p<0.0001.

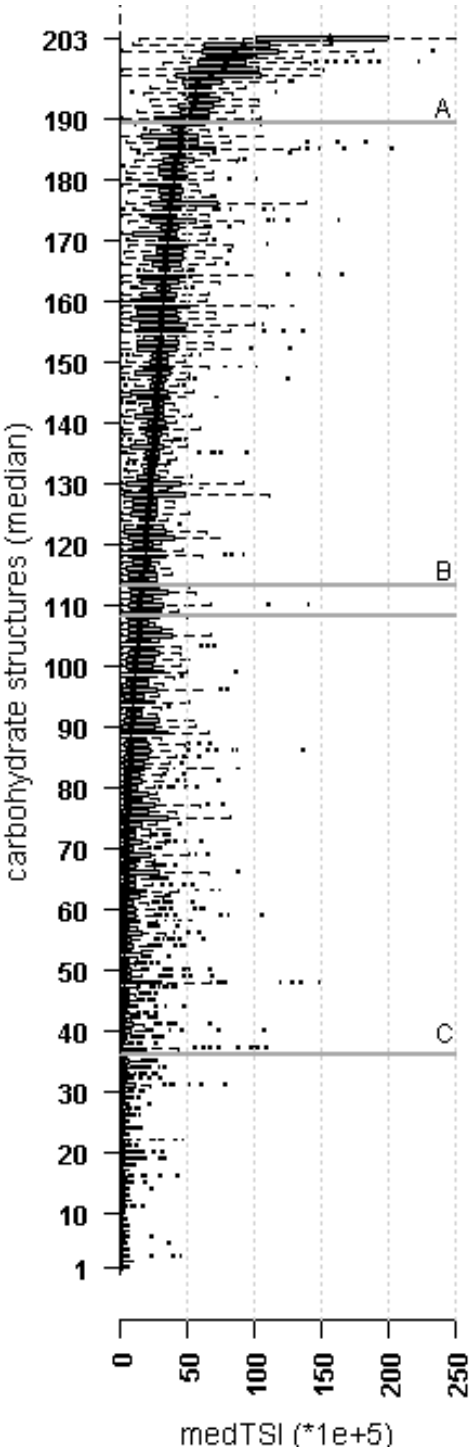


Figure 14 Median total signal intensity of anti-glycan antibody profile per glycan (n=203).
Pattern of median total signal intensities (medTSI) of AGA binding for 57 serum samples; medTSI x 10^5 on x-axis and individual carbohydrate structures sorted by median on y-axis. Grey lines indicate α-rhamnose (A); *Fucα1-2(GalNAcβ1-3)Galβ* (B); aminoglucitol (C).

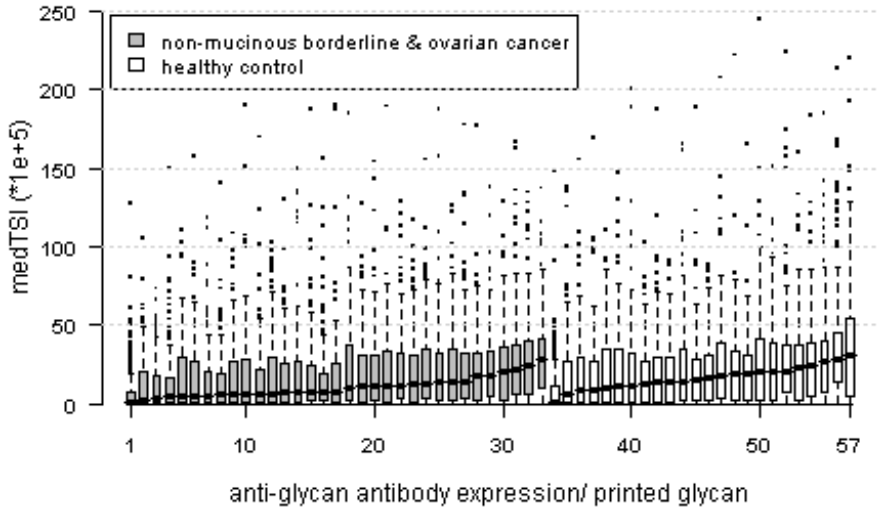


Figure 15 Anti-glycan antibody profiles for individual serum samples.
Profiles of antibody signals bound to 203 individual carbohydrate structures. Anti-glycan antibody distributions are grouped by disease status and are sorted by median per group.

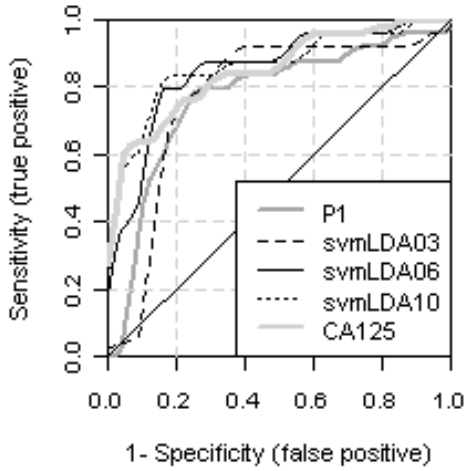


Figure 16 Receiver Operating Characteristics (ROC).
for the univariately selected top candidate P_1, standard tumor marker CA125 and multivariately selected and linearly combined glycan models. Number in svmLDA identifier indicates the number of glycans selected by support vector machine (svm) algorithm and combined in a linear discriminant model (LDA): svmLDA03, svmLDA06, svmLDA10. Carbohydrate structures involved in the linear model are described in order in Table 9.

3.2.10 Tables

	No. of patients	% of patients
Healthy control	24	42.1
Ovarian borderline tumor	2	3.5
Ovarian cancer	31	54.4
Age (yrs) (n=57)		
< 50	11	19.3
≥ 50	46	80.7
Tumor Stage (FIGO)(n=32)		
I	4	12.5
II	3	9.4
III	17	53.1
IV	8	25.0
Tumor Grade (n=30)		
G1	1	3.3
G2	9	30.0
G3	18	60.0
Histological subtype (n=33)		
Serous	19	57.6
Endometrioid	7	21.2
Clear cell	1	3.0
Transitional cell	4	12.1
Undifferentiated	2	6.1
Residual disease (n=23)		
< 1cm	12	52.2
≥ 1cm	11	47.8
Preoperative CA125 (n=53)		
< 35 U/ml	18	34.0
≥ 35 U/ml	35	66.0
Blood group (n=52)		
A	22	42.3
B	9	17.3
AB	2	3.8
0	19	36.6

Table 6 Clinicopathological characteristics of the patient cohort.
Number of patients and percentage for patient cohort incorporating age, stage, grade, histological subtype, residual disease after primary cytoreduction, preoperative standardized CA125 measurements (ELISA), and routine clinical blood group detection (ELISA) for a cohort of 57 patients.

Cluster ID	Number of glycans	Key Carbohydrate Substructure	Range of glycan length	architecture	Localization of key carbohydrate substructure	Range of pairwise correlation	Expected natural linkage
1	6	Galβ1-3GlcNAcβ1-3Galβ1-4Glcβ	4-6	*	core	0.78-0.98	GSL
9	8	$[Galβ1-4GlcNAcβ]_n$	2-6	-	varying	0.48-0.96	*N*
10	3	$[GlcNAcβ1-4]_n$	3-6	-	varying	0.91-0.96	*N*
15	3	3'-O-Su-Galβ1-4	2	-	terminal	0.69-0.84	GSL, *N*
16	5	Galβ1-4Glc or GlcNAcβ	2	-	core	0.15- 0.62	GSL, *N*
17	4	6'-O-Su-Galβ1-4GlcNAc	2	-	varying	0.43-0.81	*N*
26	4	4'-O-Su-Galβ1-4GlcNAc	2	-	varying	0.78-0.98	*N*
28	2	$[Glcα-]_n$	3,4	-	varying	0.94	GSL
30	3	Galα1-3(Fucα1-2)Galβ	3-4	-	varying	0.87-0.97	GSL, *N*
31	3	GalNAcα1-3(Fucα1-2)Galβ	3-4	-	varying	0.59-0.92	GSL, *N*
32	2	Neu5Gcβ or Neu5Acβ 2-6 Galβ1-4GlcNAcβ	3	-	core	0.87	*N*
33	4	Galβ1-3GalNAcα-	2-4	*	varying	0.66-0.86	*O*
34	8	Fucα1-3, Fucβ1-3 or Galβ1-4GlcNAcβ	1-4	*	core	0.44-0.90	*N*
35	2	Neu5Acβ2-6GalNAcα	2,4	-	varying	0.67	*O*
36	2	Galα1-4Galβ1-4GlcNAcβ	3,4	*	varying	0.64	*N*
37	2	Galβ1-4GlcNAcβ1-3GalNAcα	3,5	*	varying	0.82	*O*
38	4	(6,6'Su_2; Neu5Acα2-6; Galα1-3; Galα1-4) to Galβ1-4GlcNAcβ	2-4	*	varying	0.25-0.63	*N*
40	3	Fucα/Neu5Acα1/2-2/3/4-Galβ1-3GlcNAcβ	3	*	core	0.46-0.64	*N*

42	2	Neu5Acα2-3Galβ1-4 GlcNAcβ	3	**-**	varying	0.70	*N*
44	2	GalNAcα	4	**+**	varying	0.96	*N, O*
47	2	GlcNAcβ1-6Galβ1-3GlcNAcβ	3,4	**-**	varying	0.69	*N*

Table 7 Selected clusters with structural similarities.
Cluster identifiers, number of glycans within each cluster, core structure of cluster, glycan length (number of monosaccharides in cluster) are listed for structurally similar glycan clusters. Glycan architecture occurs as a linear tree (-), branched tree (+) or linear/ branched together (*). Localization of core carbohydrate substructures describes where the common sub-structure occurs; at core (directly after the spacer), at terminal part, or varying location. Pearson correlation coefficients for glycans are summarized by within-cluster range of values. Expected natural glycan linkage is described as Glycosphingolipid (GSL); *N*-linked (N); *O*-linked glycan (O).

	A) "Best cut-off point"		**B) Precision- recall break even point**	
	Sensitivity (%)	**Specificity (%)**	**Sensitivity (%)**	**Specificity (%)**
CA125	64.0	95.6	76.0	73.9
P_1	79.2	75.7	70.8	78.8
Chitobiose-Asn	66.7	84.4	66.7	75.0
GNb6'LN	66.7	84.8	66.7	75.7
GlcNAcβ1-6'LN	66.7	66.7	66.7	54.2
P^k	62.5	75.7	62.5	72.7
svmLDA03	87.5	69.7	75.0	81.8
svmLDA06	79.2	84.8	79.2	84.8
svmLDA10	83.3	84.8	79.2	84.8
LDA/P_1-CA125	82.6	72.0	76.0	73.9

Table 8 Discriminative power of univariate/multivariate selected candidates.
Sensitivity and specificity for "best cut-off point" and "precision-recall break even point" calculated for tumor marker CA125, individual discriminative anti-glycan antibodies and anti-glycan antibody panels for 3 (svmLDA03), 6 (svmLDA06) and 10 (svmLDA10) glycans; LN = *Galβ1-4GlcNAc*.

Rank	Carbohydrate Structure
1.	Galα1-4**Galβ1-4GlcNAcβ**
2.	Neu5Acα2-3**Galβ1-4**-(6-Su)**GlcNAcβ**
3.	Neu5Acβ2-6GalNAcα
4.	Neu5Acβ2-6**Galβ1-4GlcNAcβ**
5.	Neu5Acα2-3Galβ1-3-(6-Su)GalNAcα
6.	**Galβ1-4GlcNAcβ**1-6**Galβ1-4GlcNAcβ**
7.	**Galβ1-4**(6-O-Su)**GlcNAcβ**
8.	Galα1-3**Galβ1-4**(Fucα1-3)**GlcNAcβ**
9.	Galα1-3(Galα1-4)**Galβ1-4GlcNAcβ**
10.	GlcNAcβ1-3**Galβ1-4GlcNAcβ**

Table 9 Multivariate selected discriminating glycan structures.
Glycans ranked by their stability after 700 independent runs

3.3 Comparison of three glycan-based immunoassays: suspension array, ELISA and printed glycan array in the detection of human anti-glycan antibodies.

Manuscript in preparation

Jacob F.[1], Pochechueva T.[1], Goldstein D.R.[2,3], Huflejt M.E.[4], Chinarev A.[5], Spengler, M.[1], Fink D.[6], Bovin N.V.[5], Heinzelmann-Schwarz V.[1]

[1] *Translational Research Group, University Hospital Zurich, Switzerland;*

[2] *Institut de Mathématiques, Ecole Polytéchnique Fédérale, Lausanne, Switzerland;*

[3] *Swiss Institute of Bioinformatics, Lausanne, Switzerland;*

[4] *Cellexicon, Inc, La Jolla, CA, USA; Current address: New York University School of Medicine, Department of Cardiothoracic Surgery, Division of Thoracic Surgery and Thoracic Oncology, New York, NY, USA*

[5] *Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow 117997, Russian Federation;*

[6] *Department of Gynecology, University Hospital Zurich, Switzerland*

Correspondence to Viola Heinzelmann-Schwarz, Translational Research Group, Department of Gynecology, University Hospital Zurich, Nord I D222, Frauenklinikstrasse 10, 8091 Zurich, Switzerland; e-mail: viola@heinzelmann.ch, Telephone: +41442555374, Telefax: +41442554553

3.3.1 Abstract

Anti-glycan antibodies incorporate a vast and yet insufficiently investigated subpopulation of naturally occurring and adaptive antibodies in humans. Glycan-specific antibodies and their glycan ligands regulate various important biological processes, including cancer. The role of glycan-antibody interactions in cancer progression is still poorly understood. Recently, a variety of glycan-based microarrays were generated, allowing high-throughput profiling of the entire repertoire of glycans and glycoconjugates within one organism. There is, however, no platform established which would allow for comparison and evaluation of glycan-microarray generated candidates. In the present study we compare the three currently available 'glycan-based immunoassays', namely printed glycan array, fluorescent microsphere-based suspension array and ELISA. Our aim was to investigate glycan-based immunoassays for their efficacy and selectivity in profiling anti-glycan antibodies in a cohort of 31 patients with and without gynecological cancer. For estimation of sensitivity and specificity in the detection of anti-glycan antibodies, the AB0 blood group glycan antigens were used as ligands. All three methods reflected the known blood groups serology, as confirmed by Bland-Altman plot and receiver operating characteristics. From moderate to high concordance correlations coefficient were detected for A/B trisaccharides within all assays (CCC 0.41-0.83). A previously identified glycan candidate, which potentially allows ovarian cancer detection, P_1 trisaccharide *(Galα1-4Galβ1-4GlcNAcβ)*, was explored within all three assays. In contrast to the observations on AB0-blood group determinants, anti-P_1 antibody binding profiles displayed a variety of weak but positive results. Our results demonstrate that each assay presents unique characteristic features, presumably reflecting the difference in glycan presentation, assay conditions and detection technique. This indicates that an individual assay has to be used depending on the aspect of interest within a particular glycan-antibody interaction.

3.3.2 Introduction

Interactions of glycans, complex oligosaccharides attached to a protein or lipid, with glycan-specific antibodies mediate many important biological processes, such as pathogen recognition (1), malignant transformation (2), autoimmune disease (3), neurological disorders (4) and host-*versus*-graft rejection (5). Currently, a vast subpopulation of naturally occurring and adaptive anti-glycan antibodies are under investigation. We have previously performed

experiments identifying antibody answers to a new ovarian cancer marker, P_1, which can reflect a variety of antibody binding as well as non-specific binding.

In recent years, a number of glycan based microarray platforms, based on specific (covalent or non-covalent) immobilization of chemically conjugated glycans on a chemically modified surface, have been developed for high-throughput investigations of the glycome (6). Glycan-based arrays vary in glycan presentation, origin, assay conditions, detection methods and immobilization on flat surface or microspheres, contributing ultimately to the affinity and selectivity of glycan-binding proteins (7, 8). There is a crucial need for standardization and comparison between various forms of glycan microarrays, especially taking into account their proven scientific potential for translational research and subsequent clinical applications. The increasing number of high-throughput glycan-array studies requires the application of different bioinformatical analyses for validation and comparison of datasets. However, to date, only a few reports compared glycan microarrays and conventional ELISA, most of them only in terms of sample size and detection limitations (2, 9, 10).

In the present study we compared three independent custom-developed glycan-immunoassays in their potential and efficacy for the detection of anti-glycan antibodies in human serum; (1) printed glycan array; (2) multiplex flow cytometric suspension assay (SA); and (3) ELISA, the most established method. Printed Glycan Array (PGA) as one of the newest high-throughput microarray technologies (11, 12) allows the automated detection of an unlimited number of natural and synthesized glycans in one experimental setting. It is characterized by high sensitivity and a significant reduction of reagent consumption. Multiplex flow cytometric suspension assay (SA) (13), incorporating fluorescent microspheres with distinct spectral addresses as a solid support for ligands, was recently developed for profiling of anti-glycan antibodies (14). This type of assay combines the flexibility of ELISA with the simultaneous detection of multiple ligands in one sample with minimal reagent consumption. Standard ELISA (9, 15-17), allows screening of a limited number of ligands, is well established and widely used. Naturally occurring or chemically synthesized glycans are displayed in various ways throughout the three technologies of 'glycan-immunoassays'. In both the PGA and ELISA glycans are positioned on a flat surface. On PGA glycans are covalently attached to the surface of a glass slide in a monovalent form as ω-alkyl glycosides (Glyc-sp-NH_2), in ELISA they are physically absorbed to polystyrene as conjugates with a poly[N-(2-hydroxyethyl)acryl amide] carrier (Glyc-PAA). In the case of SA, glycans are coupled as end-biotinylated glycopolymers (18) *via* streptavidin-biotin reaction to pre-modified fluorescent beads [14]. All three glycan immunoassays display various detection specificities

resulting in certain advantages and limitations (Table 10). ELISA is more relevant for the investigation of a limited panel of glycans, PGA allows broad glycan library screening, and SA is ideal for the rapid multiplex detection of up to several dozens of samples. Therefore, all three glycan-assays could be used to study different aspects of glycan-antibody interactions.

We performed a comparative analysis of all three 'glycan-immunoassays' within a cohort of 31 healthy and gynecological cancer serum samples. As serologically investigated targets for serum anti-glycan antibodies we used blood group antigens A/B trisaccharides as biological controls (12, 19-21) as well as the candidate P_1 (*Galα1-4Galβ1-4GlcNAcβ*) which showed the highest discriminative power between a cohort of non-mucinous epithelial ovarian cancer patients and healthy controls.

3.3.3 Material and Methods

Clinical cohort

Serum samples were prospectively collected from 31 women at the Department of Gynaecology, University Hospital Zurich, after written informed consent was given. Ethical approval for this study was granted by the appropriate Ethical Board in 2006 (SPUK, Canton of Zurich, Switzerland). Two venous blood samples (12ml) were collected per patient pre-operatively in EDTA blood tubes (BD Vacutainer®, 0.184M EDTA, BD Diagnostics, Franklin Lakes, US) and stored on ice until further processing. Blood samples were centrifuged at 3000*xg* at 4°C for 10 minutes, and aliquots of the supernatant plasma frozen at –80 °C. All collected serum samples were processed within the same protocol within 3 hours.

Coupling of Suspension Array

Biotinylated glycopolymers were coupled with fluorescence labeled carboxylated beads of 5.5 µm diameter with #25, 35 and 45 spectral "addresses" (bead regions) (Bio-Rad Laboratories). The stock of beads (1.25×10^7 beads/ml) was vortexed for 30 sec and sonicated in a water bath sonicator for 30 sec before use. The bead suspension (100 µl; 1.25×10^6 beads, 0.2 nmol –COOH groups in total, according to supplier's information) was centrifuged for 4 min, 14,000 g at RT. The pellet was resuspended in washing buffer (100 µl; Bio-Plex amine coupling kit, Bio-Rad Laboratories) by vortexing and sonication and washed by centrifugation. The pellet was resuspended in 80 µl of microsphere activation buffer (Bio-Plex amine coupling kit, Bio-Rad Laboratories) by vortexing and sonication. The S-NHS solution (50 mg/ml Sulfo-N-hydroxysuccinimide sodium salt and the EDC solution (50 mg/ml 1-ethyl-3-[3,3-dimethylaminopropyl] carbodiimide hydrochloride (Pierce Biotechnology) were prepared just before use, then 10 ul of each solution were added to the microsphere suspension and vortexed

for 30 sec. The beads were incubated on a vertical rotor in the dark for 20 min at RT and subsequently centrifuged. The pellet was washed once with 500 µl of activation buffer and resuspended in 150 µl Biotin-solution (0.1 M $NaHCO_3$, pH 8.3, containing 1 µg (≈ 2 nmol) of biotin-$NH(CH_2)_6NH_2$ (Lectinity Holding, Moscow)). The suspension was rotated with medium speed in the dark for 2 h at RT. The biotinylated beads were centrifuged and resuspended in 150 µl blocking-solution (50mM Ethanolamine (Sigma-Aldrich) in 0.1 M $NaHCO_3$, pH 9.0) to block unbound sites. The beads were agitated in the dark on a rotator for 30 min at RT and centrifuged. The pellet was washed twice with 500 µl PBS (pH 7.4) and resuspended in the streptavidin-solution (400 pmol streptavidin (Bio-Rad) in 150 µl PBS per 100 µl of beads). The suspension was vortexed and agitated on a rotator in the dark for 12 h at 4°C. The beads were washed twice by centrifugation in 500 µl PBS. The Glyc-PAA-$biot_1$-solution (20 pmol biotin (conjugate of Glyc with poly[N-(2-hydroxyethyl) acrylamide in ddW) with a single end biotin group (~30 kDa, Glyc contents - 20%mol) (18)) and 1.25×10^6 streptavidin-coated beads in 150 µl PBS was added and agitated in the dark on a rotator for 6 h at RT. The modified beads were centrifuged and supernatant removed. The beads were washed twice with 500 µl storage buffer (Bio-Rad Laboratories), centrifuged and resuspended in 100 µl of storage buffer (define the buffer). The bead concentration was determined by a hemocytometer (Roth AG). The samples were stored protected from light at 4°C

Antibody binding suspension assay (SA)

A Bio-Plex Suspension Array System (Bio-Rad) which identifies and quantifies the specific reactions based on bead color and fluorescence was used for bead-based suspension assays. Data analysis was performed by Bio-Plex Manager 4.1 software. A 96-well Multiscreen HTS filter plate (Millipore) were rewetted for 5 min with 100 µl of antibody diluent (PBS-0.05 M Tris, pH 7.2, 0.25% BSA, Sigma-Aldrich). Antibody diluent incorporating (50 µl/well and 2000 beads) was added. The plate was washed three times with 100 µl washing-buffer (PBS-0.05 M Tris, pH 7.2) using a vacuum manifold (Bio-Rad). Human serum samples were diluted with antibody diluent 1:40 (50 µl/well) and added to the wells, shaken at 1,100 rpm for 30 sec on a microplate shaker and incubated in the dark for 1 h at RT on the shaker (200-300 rpm). The samples were washed three times with washing-buffer and incubated in the II antibody-mix (R-phycoerythrin conjugated goat anti-human Ig (25 ng/well IgM+IgG+IgA, H+L; Southern Biotechnology) on a shaker in the dark for 30 min. The samples were washed three times with washing-buffer and the beads were resuspended and diagnosed by the Bio-Plex-

Array reader. The data were acquired in by the internal software package (Bio-Plex Manager 4.1; Bio-Rad). The results were reported as median fluorescence intensities (MFI).

Direct ELISA

Immunoplates 96-well NUNC MaxiSorp (Thermo Fisher Scientific) were coated with the coating-solution (10µg glycans bound to poly[N-(2-hydroxyethyl)acryl amide] carrier (Glyc-PAA) (Lectinity Holdings, Moscow) per 1 ml 50mM Na_2CO_3/ $NaHCO_3$, pH9.6) for 12 h at 4°C. Carbohydrate-free PAA was applied as negative control. The plates were blocked with 1% (w/v) BSA (Sigma-Aldrich) in PBS for 40 min at 37°, washed three times with PBS containing 0.5% (v/v) Tween20. Blood serum samples were diluted 1:1000 in incubation buffer (PBS, 0.3% (w/v) BSA, 0.02% (v/v) Tween20) and incubated for 60 min at 37°. The plates were washed three times with PBS containing 0.5% (v/v) Tween20. Bound anti-glycan antibodies were detected by 60 min incubation at 37° in goat anti-human IgA + IgG + IgM conjugated (0.16µg/ml in incubation buffer) to long chain biotin (Pierce, Rockford), three times washed and incubated in streptavidin horse radish peroxidase conjugate (0.083µg/ml HRPO in incubation-buffer, Southern Biotechnology Associates) diluted in incubation buffer for 1h at 37°C. The samples were washed three times was before development by chromogen substrate 3,3',5,5'-tetramethylbenzidine (TMB) (Sigma-Aldrich). After 5min of incubation the peroxidase reaction was stopped by equal volume of 1M H_2SO_4. Optical density was measured by 450nm ELISA reader (Tecan Spectrafluor Plus, Tecan). The serum samples were set up in duplicates. For the statistical analysis their mean was used.

Printed glycan array (PGA)

Printed glycan array slide fabrication and high-throughput profiling was performed as previously described (11, 12). Briefly, monomeric glycans of 95-98% purity (Lectinity Holdings, Moscow) were diluted in 300mM phosphate buffer pH 8.5 containing 0.005% Tween 20 and printed by robotic pin deposition on *N*-hydroxysuccinimide activated glass slides (Nexterion Slide H, Schott, Jena, Germany). Glycans were printed in 50µM concentration in eight replicates. Free *N*-hydroxysuccinimide activated groups were blocked with 50mM ethanolamine in 50mM borate buffer at a final pH of 9.2. Slides were then rinsed with deionized water, dried and stored at room temperature in a desiccator. Each serum sample was diluted 1:15 in PBS (0.1% v/v Tween20 and 3% w/v BSA), vortexed for 15 seconds and incubated at 37°C for 15 min. Samples were transferred to the array slides and gently rocked in a sealed humidified incubator for 2h at 37°C. Unbound sample components were washed with a series of 0.1% and 0.001% Tween 20 in PBS. Antibody-solution, goat anti-human IgA + IgG + IgM conjugated to long chain biotin (Pierce Biotechnology) was diluted 1:100 in PBS with

0.1% Tween20 and 3% BSA and was added. Slides were incubated at RT in a humidified chamber for 45 min and then washed. Bound antibodies were visualized by incubating slides with fluorescent dye streptavidin-solution (Alexa Fluor^{555} Molecular Probes 1:1000 in PBS/0.1% Tween20) at RT for 30 min. Fluorescence signals corresponding to glycan-bound antibodies were measured and quantified using ImaGene analysis software version 6.1 and 7.5 (BioDiscovery). Signals were measured as total signal intensity (medTSI) per glycan and were expressed as median across eight within-array replicates.

Statistical Analysis

Combined graphical and statistical interpretations of method-comparison studies were performed (22) and included scatter plots combined with correlation and regression analysis (23). Difference Bland-Altman plots were combined with calculation of the 2s limits of the differences between individual methods (24, 25). To solve the problem of different data values, all data sets were standardized as follows: $z = (x_i - x_{mean}) / x_{sd.}$ Standardized data (z) were generated for each vector data set (x_i) by subtraction by their mean (x_{mean}) and division of their standard deviation (x_{sd}), and were called standardized antibody measurements (SAM). Data analysis, including calculation of mean, median, standard deviation and coefficient of variation was performed using the open source statistical programming language R (http://CRAN.R-project.org/, version 2.8.1). Statistically significant differences were proved for each method by Wilcoxon rank sum test. Concordance correlation coefficients (CCC) (26) which evaluate the degree to which pairs of observations fall on the 45° line through the origin, were calculated and compared within all independent methods (R package epiR). Direct comparison of two methods was performed using linear regression (27, 28). Median signals among all known blood groups were measured using the Kruskal-Wallis rank sum test (R package stat). All *p*-values < 0.05 were taken as significant and the tests were performed on the 5% significance level. For comparison of high *versus* low anti-glycan antibody levels in correlation with known blood groups, sensitivity, specificity (R package ROCR (29)) and area under the curve (AUC) were calculated for each method. The best cut-off between observed false negative and false positive rates was determined as "precision-recall break even point" (29), the point at which precision equals recall and predictions are made due to the prevalence within the given data.

3.3.4 Results

Controls and validation procedures used in glycan-based immunoassays.

Alpha-rhamnose was used as a "positive biological control" in PGA experiments due to the known high expression levels in healthy control individuals (30). Aminoglucitol, an opened

reduced form of *D*-glucose which is not present as a structural component of regular glycosylation, was shown to be negative for anti-glycan binding at similar values to the technical background (12). We therefore used it as a negative biological control for PGA and SA, where median fluorescence intensities (MFI) signals for aminoglucitol binding did not exceed 200 MFI and were close to the technical cut-off in all individual samples (data not shown). An empty polymer structure (PAA) was used in ELISA as a control for non-specific binding and displayed in all cases lower signals than the lowest signals for Glyc-PAA coating (data not shown).

The effectiveness of glycopolymer coupling to the fluorescent microspheres in SA was tested by human anti-glycan antibodies against A/B blood group antigens and P_1, purified from pooled human sera by affinity chromatography as described before (16). It both SA and ELISA, polyclonal anti-glycan antibodies bound to cognate glycans in a dose-dependent manner displayed no or very low cross-reactivity to other glycans.

Glycan-based immunoassay examination within AB0 blood group antigens

Antibody profiles for A_{tri} (*GalNAcα1-3 (Fucα1-2)Galβ*) and B_{tri} (*Galα1-3(Fucα1-2)Galβ*) blood group antigens were performed using SA, ELISA and PGA on 31 individual blood serum samples. Raw data were acquired for all technologies in different units; values varied from 74 to 13595 (MFI, SA), from 0.152 to 2.305 (OD, 450nm, ELISA) and from 0 to $51.28*10^5$ (medTSI, PGA). Anti- A/B antibody levels were calculated as distribution of median of 31 serum samples and additional median absolute deviation. Using **PGA**, we observed generally higher antibody binding levels to B_{tri} than to A_{tri} (p=0.0358), with higher individual variability of the latter, consistent with literature (12). The same pattern was observed in **ELISA** (p=0.0406), followed by lower individual variability of both anti-glycan antibodies, indicated by median absolute deviation.

In contrast to **PGA and ELISA**, data generation by SA revealed no significant difference between general levels of anti-A and anti-B antibody binding, accompanied by higher individual variability in anti-A antibody binding.

The appropriate distribution of the data of three immunoassays in the same units and on the same scale was achieved by standardization. PGA data were pre-processed (12) and SA skewed distributed raw data were log-transformed to improve interpretability and visualization. Next, SA, ELISA and PGA data were standardized as described above (Materials and Methods). Standardized data were further used for all statistical models, presented hereafter.

Closest relationship between SA and PGA in detection of anti A/B antibodies

Direct binary comparisons of three methods, based on scatter plots and linear regression model revealed a pattern of positive correlations, ranged from moderate to high (Figure 17). Moderate CCCs were observed for **SA and ELISA** in A_{tri} and B_{tri} antigen (CCC 0.68 and 0.71, respectively, Table 11). These positive correlations were displayed by the proximity of linear regression lines to the equality line (Figure 17) for both antigens (A_{tri} and B_{tri}, Table 11). The highest correlation was observed between **SA and PGA** (Figure 17), with CCC 0.83 for A_{tri} and 0.70 for B_{tri} (Table 11). The weakest correlation was observed between **ELISA and PGA** (Figure 17), with a CCC of 0.54 (A_{tri}) and 0.41 (B_{tri}) (Table 11). Observed correlation patterns for all three comparisons validated by F-statistics showed significant p-values ($p<0.05$). This supported the hypothesis that anti-glycan antibody levels across the three methods correlated to each other. No significant difference was found for both linear regression models in comparison of A_{tri} *versus* B_{tri} ($p> 0.05$) in all three binary comparisons, which indicated subject (here: glycan)-independent observations.

Correlation characteristics for A antigen were comparably more distinctive than for B antigen (in SA *versus* PGA and PGA *versus* ELISA, Table 11)

Additionally, the correlations between methods were examined (Figure 18) and proved glycan independent observations. It also demonstrated similar tendencies within both **SA and PGA**, showing the smallest absolute standard deviation (2*s.d.) for A_{tri} (1.16SAM) and a bigger absolute standard deviation for B_{tri} (1.54SAM) (Figure 18). The most pronounced difference was observed between **PGA and ELISA** with 1.91SAM for A_{tri} and 2.16 SAM for B_{tri} (Figure 18). None of the immunoassays showed any difference in the detection of anti-A/B antibodies, although A_{tri} had a more distinctive distribution pattern compared to B_{tri}.

All immunoassays reflect adequately the standard serological distribution of AB0 blood types

The well known AB0 blood group carbohydrate antigens were used for measurement of assay capacity in the detection of anti-glycan antibodies. AB0-specific antibodies (isohaemagglutinins) belong to a group of naturally occurring antibodies (31) which can be determined by haemagglutination test, ELISA (15, 17) and FACS (32). To define how applied assays distinguish between blood groups we performed non-parametric tests for all three data sets. All ‘glycan-based immunoassays’ were able to distinguish blood groups (A (n=13), AB (n=2), B (n=6) and O (n=10)) adequately (Figure 19). As expected, highly significant differences ($p<0.01$) between known blood group antigens A_{tri} and B_{tri} were found in each applied method (SA, ELISA and PGA). Strong discrimination of antibody levels for A_{tri} was observed in all three methods with highest in SA ($p<0.0001$). Less significant discriminators were achieved for antigen B_{tri} (SA, $p=0.0058$; PGA, $p=0.0065$). Distribution of anti-A/B

antibodies using a pool of immunoglobulin subtypes (IgG, IgA, IgM) determined by all three immunoassays were concordant with classical haemagglutination data (33).

High diagnostic performance in all glycan-based immunoassays

The accuracy of applied assays to discriminate between expected anti-glycan antibody levels was evaluated by ROC. Data from blood groups with low anti-A_{tri} antibody levels (BG A and AB) were combined, as were the data from BG with higher anti-A antibody levels (BG B and O). SA revealed the best discrimination character between low and high anti-A antibody levels with excellent selectiveness (sensitivity 93.3%, specificity 93.7%, AUC 0.96). Equal AUC values were achieved by ELISA (0.95) and PGA (0.93) (Figure 20).

Interestingly, the overall pattern changed for anti-B antibody levels and sensitivity, specificity and AUC values decreased in PGA, SA and ELISA (AUC 0.85, 0.87 and 0.78, respectively) (Figure 20).

Glycan-based immunoassays showed diverse expression for P_1 (Galα1-4Galβ1-4GlcNAcβ)

Based on the findings that anti-A/B antibody profiles represent an assay independent pattern we further investigated potential ovarian cancer biomarker, antigen P_1 (*Galα1-4Galβ1-4GlcNAcβ*). 'Glycan-based immunoassays' were plotted against each other and their relationship determined by linear regression analysis (Figure 21). In comparison to the A/B blood group-based assays, anti-P_1 correlations across the three assays appeared to be much weaker. Best relationships with moderate correlation (CCC 0.49) were detected between SA and ELISA assays. Almost no relationship could be observed comparing of PGA *versus* ELISA and PGA *versus* SA (CCC 0.35 and 0.27, respectively), which is consistent with the findings that SA and ELISA do not generate the same results as PGA. Within cancer samples, the CCC as well as the linear regression line was inverted in the comparisons PGA *versus* ELISA or SA (Figure 21). Within the healthy controls the correlation coefficient revealed at least moderate or low correlations (CCC 0.52 (SA *versus* PGA) and 0.29 (PGA *versus* ELISA)). Anti- P_1 antibody levels between healthy and malignant cases showed only significant discrimination character in the PGA experiment (p=0.01) (Figure 22).

3.3.5 Discussion

In the present study we statistically measured and compared the efficacy and selectivity of three custom-developed immunoassays in profiling of anti-glycan antibodies in individual human serum samples. Two of the investigated glycans belonged to the subpopulation of AB0 blood group antigens and one was previously described as potential ovarian cancer biomarker.

With the use of different statistical models we have shown, that all glycan immunoassay results for anti-A/B antibodies were positively correlated, from moderate to high correlations, and demonstrated the same tendencies for both blood group antigens independent of method used. For all methods the correlation patterns for A antigen were comparably more distinctive than for B antigen.

All three methods adequately reflected known serological isohaemagglutinin distribution within four blood groups and showed high diagnostic performance in discrimination of blood groups, but best for SA and PGA.

However, in comparison to the A/B blood group-based assays, anti-P_1 correlations across the three assays considerably changed. PGA demonstrated a discrimination power between cancer and healthy serum samples whereas the other two methods did not result in reproducible results. This could reflect a true cancer association of P_1 antibody levels, but because of the cohort size strong conclusions cannot be drawn. Differences in the various assay conditions like different concentrations of glycans as well as serum dilutions (Table 10) have to be drawn into consideration. Whilst this finding is of interest, our main aim in this study was to compare three 'glycan-immunoassays' on their expression of well known AB0 blood group anti-glycan antibodies. Our results indicate a stronger relationship among these assays were much stronger in case of AB0 blood group antigens than in case of P_1. This could be explained by their strong biological importance, or high affinity for anti- A/B antibodies (21). Although the affinity and biological importance of anti-P_1 antibodies was not investigated, they might comprise rather a subpopulation of anti-glycan antibodies with lower affinity (and/or specificity). As in the case of PGA glycans are presented as monomers in a highly organized manner with limited flexibility (and possibly restricted accessibility) (Table 10), presumably only a subpopulation of anti-P_1 specific antibodies but not all the spectrum of possible P_1 binders was detected. In contrast to PGA, SA and ELISA are assays based on polymeric glycan presentation which allows multivalent binding between glycans and glycan directed antibodies (Table 10). We assume that SA and ELISA reflect a broader spectrum of glycan-antibody interactions. SA provides more degree of freedom than in PGA as flow conditions and polymeric glycan presentation contribute to the binding and interaction (Table 10) (13, 14). As a result, ELISA and SA might be more susceptible than PGA to non-specific binding, like the direct binding of heterophilic serum antibodies. Binding of heterophilic antibodies (rheumatoid factor IgMs) was previously reported for SA and ELISA, could interfere data readout causing false positive or negative results (34, 35) On the other hand, assay based on multivalent and

polymeric ligand presentation seem therefore to be closer to natural or *in vivo* conditions, which was shown by a good discrimination power for anti-A/B antibodies.

We conclude that for biomarker research the use of only one assay with appropriate sensitivity, specificity and reproducibility might be sufficient to identify a specific disease. In contrast, it is not relevant to draw crucial conclusions based on data from one assay in the case of basic research, as each glycan, showing important physiological patterns in one assay, should be confirmed by other approaches. Each method is specific to describe glycan binding patterns (12), glycan binding partners (36) or motif-based analysis of specify glycan-binding proteins (37). For investigations into glycan-protein interactions the cumulative approach based on the use of several methods for better data interpretation seems more preferable, similar to cDNA or oligonucleotide microarrays being validated using RT-PCR as well as Northern blotting on the basis of gene expression levels.

Based on our findings, the usefulness of each glycan-immunoassay strongly depends on the task, taking into account the advantages and limitations of each method. SA is better than ELISA in consuming a lower amount of glycopolymers (picomolar *versus* micromolar amounts per experiment), whilst SA is more suitable than PGA due to a greater range of flexibility and assay reconfiguration. On the other side, PGA is exceptionally convenient for high-throughput screening of big glycan libraries whereas ELISA is restricted in this aspect. Thus, an adequate usage of all discussed glycan-based immunoassays broadens the methodological flexibility in the area of glycobiology.

3.3.6 Acknowledgments

This work was supported by the Cancer League of the Canton of Zurich (V.H.S.); Oncosuisse (02115-08-2007 to V.H.S.); Swiss National Foundation (320000-12543 to V.H.S.), European Science Foundation (V.H.S.); Cellexicon, Inc. La Jolla, CA, USA; National Cancer Institute, USA, grant 1U01CA128526 (to M.E.H).

3.3.7 References

1. de Boer AR, Hokke CH, Deelder AM, & Wuhrer M (2008) Serum antibody screening by surface plasmon resonance using a natural glycan microarray. *Glycoconj J* 25(1):75-84.
2. Wang CC*, et al.* (2008) Glycan microarray of Globo H and related structures for quantitative analysis of breast cancer. *Proc Natl Acad Sci U S A* 105(33):11661-11666.
3. Dotan I*, et al.* (2006) Antibodies against laminaribioside and chitobioside are novel serologic markers in Crohn's disease. *Gastroenterology* 131(2):366-378.
4. Brettschneider J*, et al.* (2009) Serum anti-GAGA4 IgM antibodies differentiate relapsing remitting and secondary progressive multiple sclerosis from primary

progressive multiple sclerosis and other neurological diseases. *J Neuroimmunol* 217(1-2):95-101.
5. Galili U (1999) Significance of anti-Gal IgG in chronic xenograft rejection. *Transplant Proc* 31(1-2):940-941.
6. Culf AS, Cuperlovic-Culf M, & Ouellette RJ (2006) Carbohydrate microarrays: survey of fabrication techniques. *Omics* 10(3):289-310.
7. Lewallen DM, Siler D, & Iyer SS (2009) Factors affecting protein-glycan specificity: effect of spacers and incubation time. *Chembiochem* 10(9):1486-1489.
8. Oyelaran O, Li Q, Farnsworth D, & Gildersleeve JC (2009) Microarrays with varying carbohydrate density reveal distinct subpopulations of serum antibodies. *J Proteome Res* 8(7):3529-3538.
9. Galanina OE, Mecklenburg M, Nifantiev NE, Pazynina GV, & Bovin NV (2003) GlycoChip: multiarray for the study of carbohydrate-binding proteins. *Lab Chip* 3(4):260-265.
10. Willats WG, Rasmussen SE, Kristensen T, Mikkelsen JD, & Knox JP (2002) Sugar-coated microarrays: a novel slide surface for the high-throughput analysis of glycans. *Proteomics* 2(12):1666-1671.
11. Blixt O*, et al.* (2004) Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. *Proc Natl Acad Sci U S A* 101(49):17033-17038.
12. Huflejt ME*, et al.* (2009) Anti-carbohydrate antibodies of normal sera: findings, surprises and challenges. *Mol Immunol* 46(15):3037-3049.
13. Nolan JP & Sklar LA (2002) Suspension array technology: evolution of the flat-array paradigm. *Trends Biotechnol* 20(1):9-12.
14. Bovin NV & Huflejt ME (2008) Unlimited glycochip. *Trends Glycosci. Glycotechnol.* 20(115):245-258.
15. Buchs JP & Nydegger UE (1989) Development of an ABO-ELISA for the quantitation of human blood group anti-A and anti-B IgM and IgG antibodies. *J Immunol Methods* 118(1):37-46.
16. Obukhova P, Rieben R, & Bovin N (2007) Normal human serum contains high levels of anti-Gal alpha 1-4GlcNAc antibodies. *Xenotransplantation* 14(6):627-635.
17. Rieben R, Buchs JP, Fluckiger E, & Nydegger UE (1991) Antibodies to histo-blood group substances A and B: agglutination titers, Ig class, and IgG subclasses in healthy persons of different age categories. *Transfusion* 31(7):607-615.
18. Chinarev AA, Galanina OE, & Bovin NV (2010) Biotinylated multivalent glycoconjugates for surface coating. *Methods Mol Biol* 600:67-78.
19. Oyelaran O, McShane LM, Dodd L, & Gildersleeve JC (2009) Profiling Human Serum Antibodies with a Carbohydrate Antigen Microarray. *J Proteome Res*.
20. Yamamoto F (2004) Review: ABO blood group system--ABH oligosaccharide antigens, anti-A and anti-B, A and B glycosyltransferases, and ABO genes. *Immunohematology* 20(1):3-22.
21. Watkins WM (2001) The ABO blood group system: historical background. *Transfus Med* 11(4):243-265.
22. Standards NCfCL (1995) Method comparison and bias estimation using patient samples, approved guideline. Villanova, PA: NCCLS publication EP9-A.
23. Westgard JO & Hunt MR (1973) Use and interpretation of common statistical tests in method-comparison studies. *Clin Chem* 19(1):49-57.
24. Altman DG & Bland JM (1983) Measurement in medicine: the analysis of method comparison studies. *Statistician* 32:307-317.
25. Bland JM & Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1(8476):307-310.

26. Lin LI (1989) A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 45(1):255-268.
27. Wilkinson GN & Rogers C (1973) Symbolic descriptions of factorial models for analysis of variance. *Applied Statistics* 22:392–399.
28. Chambers J (1992) *Linear models* (Wadsworth & Brooks/Cole).
29. Sing T, Sander O, Beerenwinkel N, & Lengauer T (2005) ROCR: visualizing classifier performance in R. *Bioinformatics* 21(20):3940-3941.
30. Dotan N, Altstock RT, Schwarz M, & Dukler A (2006) Anti-glycan antibodies as biomarkers for diagnosis and prognosis. *Lupus* 15(7):442-450.
31. Landsteiner K (1901) Ueber Agglutinationserscheinungen normalen menschlichen Blutes. *Wiener klinische Wochenschau* 14:1132-1134.
32. Stussi G*, et al.* (2005) Isotype-specific detection of ABO blood group antibodies using a novel flow cytometric method. *Br J Haematol* 130(6):954-963.
33. Brecher ME (2002) *American Association of Blood Banks, Technical manual* (Bethesda) 14 Ed.
34. Martins TB, Augustine NH, & Hill HR (2006) Development of a multiplexed fluorescent immunoassay for the quantitation of antibody responses to group A streptococci. *J Immunol Methods* 316(1-2):97-106.
35. Waterboer T, Sehr P, & Pawlita M (2006) Suppression of non-specific binding in serological Luminex assays. *J Immunol Methods* 309(1-2):200-204.
36. Gout E*, et al.* (Carbohydrate Recognition Properties of Human Ficolins: GLYCAN ARRAY SCREENING REVEALS THE SIALIC ACID BINDING SPECIFICITY OF M-FICOLIN. *J Biol Chem* 285(9):6612-6622.
37. Porter A*, et al.* (A motif-based analysis of glycan array data to determine the specificities of glycan-binding proteins. *Glycobiology* 20(3):369-380.

3.3.8 Figures

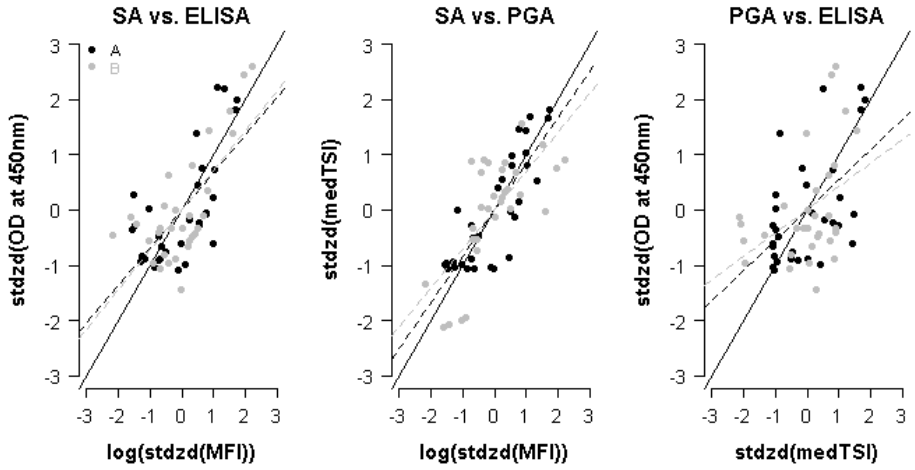


Figure 17 Comparison of 'glycan-based immunoassays' by scatter plot.
Methods were plotted against each other among all possible combinations. The *x*- and *y*- axis shows the standardized signals (see "Results") for each method. Dashed lines indicate linear regression model for blood group antigen A_{tri} (black) and blood group antigen B_{tri} (grey). Black line indicates intercept of 0 and slope of 1. Each dot is represented by two-method measurement of one serum sample (n=31).

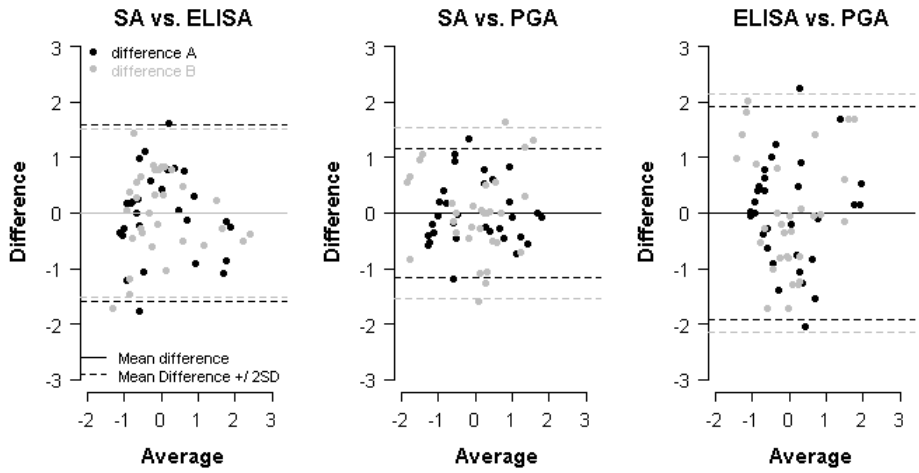


Figure 18 Binary comparison of 'glycan-based immunoassays' by Bland-Altman plot.
The averages of each two methods standardized data (*e.g.* [SA– ELISA]/2) are plotted on *x*- axis, and *y* axis refers to the absolute difference between each two methods ([SA– ELISA]). Upper and lower dashed lines indicate 95% limit of agreement for blood group antigen A (black) and blood group antigen B (grey). Black line indicates the absence of absolute difference between two methods. Each dot indicates the value of difference between two methods in each serum sample (n=31).

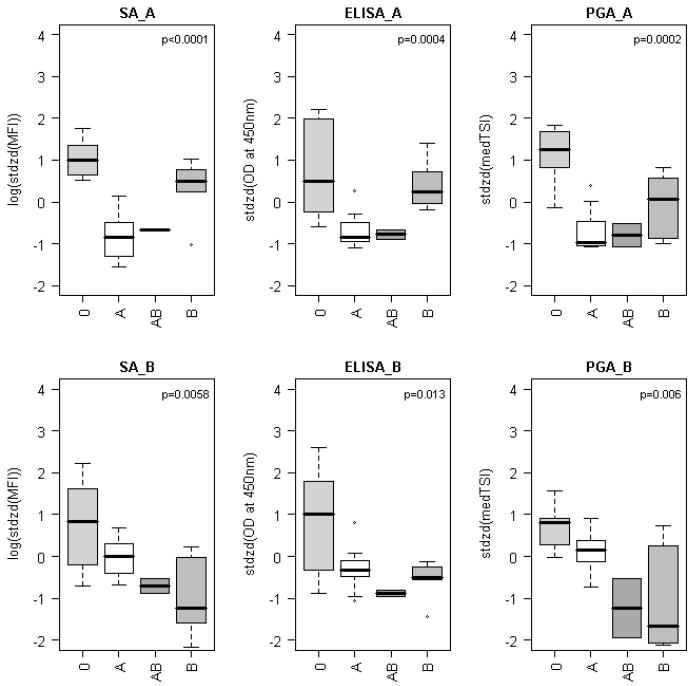


Figure 19 Binding of serum anti-A/B antibodies to A_{tri} and B_{tri} in SA, ELISA and PGA.
Boxplots demonstrate the distribution of anti-glycan antibodies directed to glycans A_{tri} (plotted as SA_A; ELISA_A; PGA_A) and B_{tri} (SA_B; ELISA_B; PGA_B) within each blood group (A, B, AB and O). The median per blood group is indicated by the horizontal black bar in a box. The bottom and the top of a box represent the lower and upper quartile, respectively. The length of a box is designated as interquartile range. The length of dotted lines, whiskers, is within 1.5 times of interquartile range. The data outside the whiskers are considered as outliers and are represented by empty circles. Blood groups are shown on x-axis. Standardized signals for each method are presented on y-axis. Kruskal-Wallis *p*-value indicates the equality of population medians among blood groups.

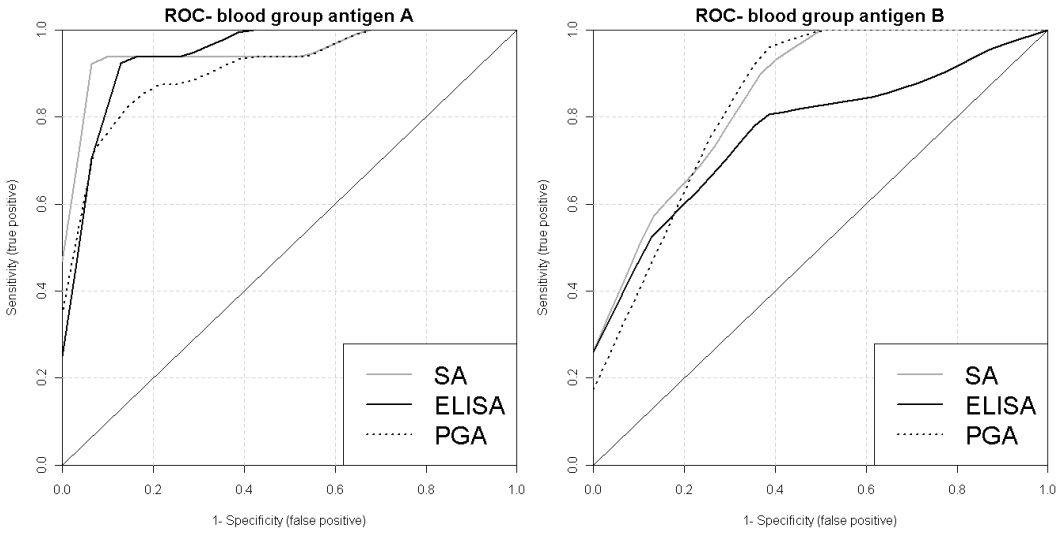


Figure 20 ROC curves for anti- A/B antibodies.
Data from BG with expected low anti-A antibody levels (BG A and AB), as well as data from BG with higher anti-A antibody levels (BG B and O), were combined to assess the binary classification test characteristics (ROC curves) for antigen A in comparison of three methods. In reverse, the data from BG B +AB *versus* BG A+O were also combined for comparison of three methods, based on antigen B.

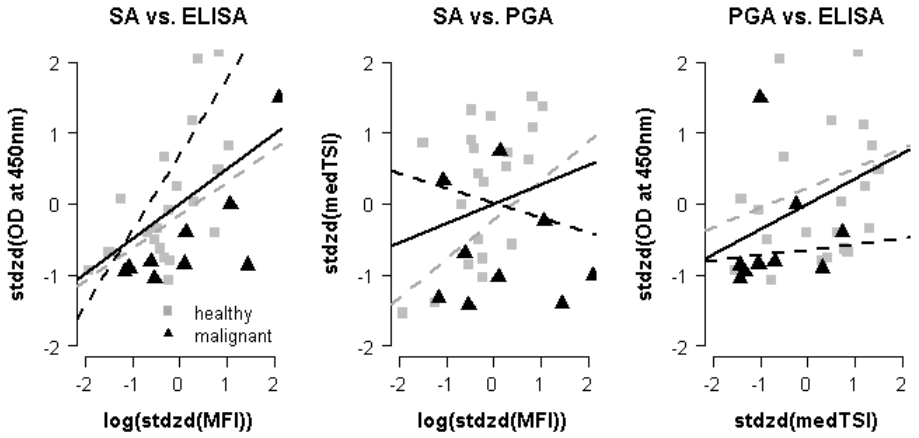


Figure 21 Scatter plots for anti-P_1 antibody levels in SA, ELISA and PGA.
Methods were plotted against each other to describe the relationship of measurements (n=31). SA *versus* ELISA revealed best assessing agreement (CCC 0.49). No correlation could be detected between SA or ELISA *versus* PGA (CCC 0.27 or 0.35, respectively). Grey quadrates represent serum samples from healthy donors and black triangles refer to samples from donors with gynecological malignancies. Linear regression is indicated by black line for all serum samples (n=31), dashed grey line for healthy samples (n=22) and black dashed line for sera of malignant cases (n=9). The *x*- and *y*- axis shows the standardized values of each method.

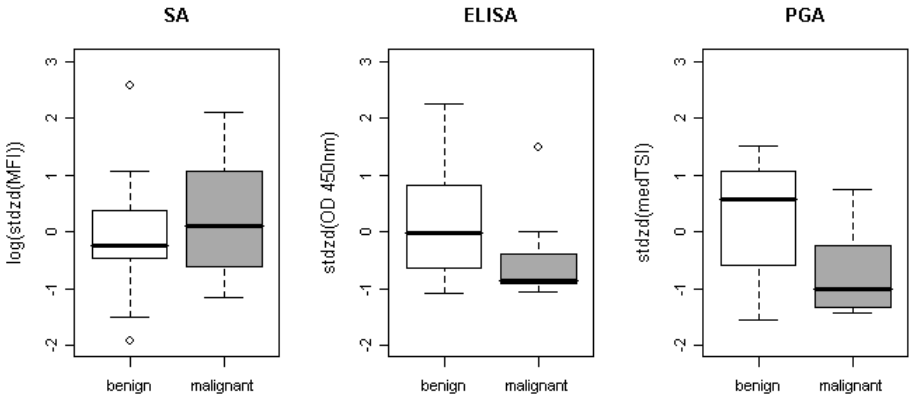
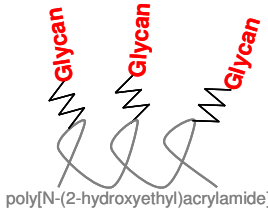
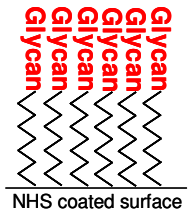
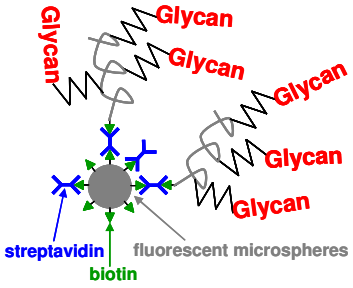


Figure 22 Binding of anti-P_1 antibodies in SA, ELISA and PGA.
Boxplots demonstrate the distribution of anti-glycan antibodies directed to P_1 measured by three immunoassays. Benign (n=22) and malignant (n=9) cases are plotted separately.

3.3.9 Tables

	ELISA	**Printed Glycan Array**	**Suspension Array**
Schematically presentation of glycans			
composition	polymeric	monomeric	polymeric
dimensional presentation	2D	2D	3D
throughput	monoplex	multiplex (2 < n)	multiplex (2-200)
platform	NUNC™ MaxiSorp™ 96 well plates	Schott NHS-activated PEG-coated Nexterion slides H	Luminex carboxylated fluorescent microspheres
secondary detection	goat anti-human IgA, IgG & IgM (H+L) long chain biotin conjugated	goat anti-human IgA, IgG & IgM (H+L) long chain biotin conjugated	goat anti-human IgA, IgG & IgM (H+L) R-phycoerythrin labeled
molecular reporting	Streptavidin Horse Reddish Peroxidase conjugated	Streptavidin Alex Fluor^{555}	n/a
dilution: blood serum	1:1'000	1:15	1:40

conc.: 1st detection	0.16 µg/ml	20 µg/ml	25 ng/well
conc.: 2nd detection	0.083 µg/ml	2 µg/ml	n/a
data acquisition	TMB as substrate; peroxidase reaction; 450nm	Fluorescent measurement at 533 nm	Fluorescent measurement at 533 nm
technical read out	optical density at 450nm	median fluorescent total signal intensities of eight replicate	median fluorescent signal intensity (within 100 beads)
assay conditions	static	static	flow

Table 10 Methodological description of three 'glycan-based immunoassays'.

	SA *versus* ELISA		**SA *versus* PGA**		**PGA *versus* ELISA**	
	A	**B**	**A**	**B**	**A**	**B**
CCC	0.68	0.71	0.83	0.70	0.54	0.41
95% CI	0.43-0.83	0.49-0.85	0.68-0.91	0.46-0.84	0.24-0.75	0.08-0.67
R-squared	0.46	0.51	0.69	0.49	0.296	0.175
F-test *p*-value	<0.001	<0.001	<0.001	<0.001	0.001	0.019

Table 11 Relationship between 'glycan-based immunoassays'.
Concordance correlation coefficient (CCC); A (A_{tri}); B (B_{tri}); printed glycan array (PGA); suspension array (SA); confidence interval (CI).

4 Discussion

4.1 High-Throughput Profiling for Epithelial Ovarian Cancer Biomarker Detection

Epithelial ovarian cancer is the fourth most common cause of cancer death in women in the Western world and the leading cause of death from gynaecological malignancies worldwide (1). Due to unspecific symptoms like bloatedness or abdominal distension, 75% of women get diagnosed in advanced stage disease with a five-year survival of 10 to 30% (2). Despite advances in surgery and chemotherapy relapse rates are as high as 50% after six cycles of standard Carboplatin/ Paclitaxel chemotherapy, with a median survival after recurrence of only two years (1, 3).

There is thus a high demand to utilize molecular expression profiles for the classification of ovarian tumors into distinct, clinically relevant subtypes and for the prediction of individual clinical outcomes by generating a personal "biosignature" (4). Transcriptional genome profiling technologies allow the analysis of the expression of thousands of genes of a specific biological system in only one experiment (5). Gene expression profiles of two RNA samples that are simultaneously hybridized onto a microarray chip can be compared and allow investigators to address important biological questions that could not be answered with traditional expression-based technologies as Northern blots, *in situ* hybridization, or real time PCR. Since the first ovarian cancer transcriptomic studies in 1999 (6-8), a rapid development of microarray technology has taken place, including the development of SEREX as the first high-throughput technology which examined protein expression (9). Meanwhile a linear increase in high-throughput profiling studies in the field of ovarian cancer has taken place (10). The immense amount of data gained through these experiments has enabled some insight into the heterogeneity of ovarian cancer and has shown some biomarkers which could evolve as new diagnostic, prognostic or therapeutic targets in the future. A summary on all high-throughput technologies published from 1999 to 2007 in the field of ovarian cancer (10) has displayed important findings but also numerous limitations which need to be overcome in future studies.

4.1.1 Current Limitations and Future Integrative Research

High-throughput expression profiling studies in the field of genomics and proteomics (proteogenomics) allow the detection of thousands of genes and proteins from very small samples of tissue, cell lines, blood or other body fluids, and have thus revolutionized our knowledge and understanding of ovarian cancer. During the last eleven years this technology

has been extensively used to unravel the ovarian cancer genome, transcriptome, and proteome, thus changing various paradigms, which were manifest for decades. A linear increase in profiling studies has left us currently with over 237 studies published between 1999 to 2007, which have used various platforms and clinical samples as well as having addressed numerous questions (10). The growing diversity of study questions during the recent years clearly demonstrates three major problems in high-throughput ovarian cancer profiling, namely the need

a) to define homogenous subtypes of ovarian cancer and its adequate normal controls,
b) to develop assays which are easy performable, reliable, comparable and reproducible,
c) to openly distribute study results.

Transcriptomic profiling, involving RNA molecules in a given cancer, is a widely used technology for the generation of gene expression data on a genomic scale with many significant results derived by several microarray platforms. Oligonucleotide arrays and cDNA microarrays are increasingly used by researchers interested in broad gene expression patterns, allowing them to compare the current activity of thousands of genes in specific tissues and cell lines under various study conditions, namely with and without genetic defect, or with and without treatment. Nowadays, the main problems regarding comparability occur due to data management and pre-processing, statistical analysis, distribution and storage of datasets.

A large degree of variability has been found in gene expression patterns investigated in tissues (48.3%), even within supposedly genetically identical individuals. Causes were identified in the time of day the experiments were performed and patient characteristics like age, physiological state, menstrual status, response to disease (10), demonstrating that the robustness of the method is not given. Additionally, the heterogeneity of the disease and its potential precursor lesions have to be defined before any biomarkers with high sensitivity and specificity can be detected. Improvements could be achieved by standardized collection and storage of specific databases. Therefore, we have developed a specific database (PEROV, Peritoneal and Ovarian Cancer data base, Microsoft Access) which allows the storage of clinical patient characteristics as well as translational research results per patient. This database links both personal and clinical parameters, biochemical measurements such as CA125 and experimental work. It therefore enables the evaluation of translational research experiments by linking antibody profiles to individual tumor characteristics.

Another well known problem in gene expression profiling is the broad spectrum of available techniques. Requirements for fabrication, cross-hybridization, target redundancies, and ultimately the use of common platforms which are reproducible and compatible have to be recommended (11). It has been shown that different public databases, *e.g.* GEO (12) and ArrayExpress (13) differ in availability and assessment of data in their raw format, so that not sufficient quality data are available to enable meta-analyses (14). In the majority of publicly available datasets the raw data were not found (14). A possible solution was suggested by Brazma *et al.* in 2001, who published a proposal for guidelines describing the minimum information about a microarray experiment (MIAME). These days most scientific journals require the deposition of microarray derived data for independent verification studies (15). Due to these requirements data are increasingly available for the reproduction of previous findings, comparison and combination of other datasets or to test new applied bioinformatical tools and relevant mathematical procedures. Microarray data have been used to build secondary databases like the *Arabidopsis* genome based GENEVESTIGATOR (16), the cancer microarray database ONCOMINE (17), or the ArrayExpress Gene Expression Atlas (13). Based on our analysis in ovarian cancer proteogenomic studies, raw data of transcriptomic based high-throughput studies (63.7%) were only in 12% available within the MIAME guidelines (10). Unfortunately, most of the remaining studies published only gene lists of limited utility for integration analysis.

Meta-analyses have emerged in the field of clinical investigations as the gold standard for the comparison and combined analysis of existing studies. It is well accepted that only meta-analysis can circumvent the problems inherent to studies with low statistical powers due to low samples sizes (18). Such analyses performed on cancer microarray data are only feasible by independent studies which investigated the same set of genes and the same clinical aim. In a meta-analysis performed by us we combined datasets of five publicly available raw data sets (19-23) and generated a list of consistently and significantly de-regulated genes which are potentially associated with ovarian cancer (unpublished data). This list included 25 most significantly up-regulated genes, including previously identified and validated ovarian cancer biomarkers like *HE4* and *DDR1*. Both have been described, serum protein (sandwich ELISA) HE4 as potential tumor marker in ovarian cancer (24-26), and tissue protein (immunohistochemistry) DDR1 as marker of early ovarian oncogenesis (27). The second most significantly up-regulated gene identified in our meta-analysis was Growth arrest-specific 6 (*GAS6*), and showed in our immunohistochemical studies an association with ovarian cancer survival (unpublished data).

The fast growing diversity of experimental designs and analytical technologies complicates the discovery and evaluation of experimental microarray data. Based on MIAME guidelines, concepts of reporting standards have been expanded to other high-throughput technologies, prevalent in proteomic experiments: a) the minimum information about proteomics experiments (MIAPE) (28), b) *in situ* hybridization and immunocytochemistry- Minimum Information Specification For *In Situ* Hybridization and Immunohistochemistry Experiments (MISFISHIE) (29, 30), c) and biomedical investigations- The Minimum Information for Biological and Biomedical Investigations (MIBBI) (31). These minimum information checklists can promote transparency in experimental reporting, enhance accessibility to data and support effective quality assessment, increasing the general value of the experimental data (31). The next era of high-throughput profiling research will rely on comparative analytical data generated from various data sets to answer questions related to causality and association (32), molecular function and regulation (33).

4.1.2 Heterogeneity of the Disease

It is becoming increasingly clear that epithelial ovarian cancer (EOC) is a heterogeneous disease. The way it is currently classified is by histopathology, however, the future might be by molecular profiling and might ultimately define the clinical outcome. Reviewing all published proteogenomic profiling based publications in EOC, we focused on the histology, stage and grade of ovarian tumors and revealed within this group various genetically distinct diseases. Different gene expression patterns separated low grade *versus* high grade SOC, early *versus* late stage EOC, EOC with local disease *versus* omental metastatic disease, or EOC with high likelihood to relapse early (10). Although high-throughput profiling technologies are promising, further improvements in technology and study design combined with pathway analysis are necessary to target innovative molecular therapy (34). A low disease incidence and substantial molecular heterogeneity among EOC subtypes hamper the speed of elucidating new insight. In consistence with recent observations (10), changes between the different histological subtypes seem to be associated with specific genetic expression patterns regardless of the substage and grade of the disease. As to tumor morphology, it should be noted that the histological origin of the different ovarian cancer subtypes is still a matter of debate.

Historically, the choice of treatment was based on pathological characteristics of the cancer, such as histology, invasion, tumor grade, size and nodal status. There is an observable trend which seems to accept that in the future cancer treatment will depend on an individual

patients molecular signatures (35). This implies a linkage between molecular diagnostics and therapeutics for the benefit of individualized therapy at each point in disease progression (36). Depending on the question addressed, the method used to screen for the molecular answer will decide the quality of the results. Whilst subtype analyses to distinguish between different genetic backgrounds might be answered using genomic profiling, diagnostic or prognostic marker detection needs to be addressed differently. Based on our analysis we conclude that genomic profiling experiments in ovarian cancer might be insufficient to investigate questions related to the etiology of ovarian cancer or to define new biomarkers, as physiological effects will depend on the relative distribution of gene products and its interaction with other genes (10, 34). As the biological effect cannot be determined on genomic or proteomic level, post-translational modification screening will facilitate the most information for clinical practice and should therefore be included into our current screening approaches (34, 37).

4.1.3 Ovarian Cancer Biomarker Research

Various potential biomarkers have been discovered for early detection as well as post-treatment monitoring using proteogenomic profiling strategies. Classical transcriptional high-throughput technologies are limited by the detection of single biomarkers instead of a group of identifiers. Some of the detected and validated biomarkers are already under clinical evaluation using specifically designed sandwich ELISA, *e.g.* the most promising candidate human epididymis protein 4 (HE4) (24-26, 38, 39) or osteopontin (OPN) (40, 41). To our knowledge, however, there is no single biomarker which is superior to the current standard tumormarker CA125. This could be caused by technical limitations and the natural complexity or individual variability of gene and protein expression signatures. The heterogeneity of the disease also influences the detection of biomarkers of high performance, thus, new strategies have been proposed to identify more significant markers of early detection.

Meta-analytical approaches were chosen to combine individual tumor markers for the benefit of improved significance and specificity. Whilst this requires complete raw datasets, we developed an automated overlapping ovarian (OLOV) database to integrate and intervalidate the results of cDNA microarray- and oligonucleotide array-based transcription profiling studies (10, 27). The gene with the highest significant number of overlaps within studies *OI4* (COL1A2, collagen, type I, alpha 2), followed by *VEGF*, a gene whose protein is targeted *via* mAb bevacizumab (Avastin®) and is currently under examination in clinical trials (42). The up-regulated bone morphogenetic protein 7 (BMP7) has been shown to be a potential inhibitor of transforming growth factor-beta (TGF-beta) which is involved in ovarian cancer signaling

(43). In ovarian cancer, heavily glycosylated MUC1 (44) was identified and its protein is known to show similar significance and specificity as CA125 and HE4 (26, 38). All of these proteins seem to be directly or indirectly linked to ovarian carcinogenesis. Other potential candidates which are overexpressed in ovarian cancer have already treatment options available, like Gleevec® (inhibit a family of tyrosine kinases) or Iressa® (targets EGFR family), which have shown a cytotoxic effect in recurrent cancers and are currently in phase two and three clinical studies in combination with or without Carboplatin and Paclitaxel as first line therapy. Gene therapies targeting *p53*, *BRCAI*, *erbB2* are already in phase I to III clinical trials (45), however, new prognostic and predictive markers have not yet shown improved sensitivity and specificity compare to the current tumor marker CA125.

Since genomic studies lack sensitivity and specificity in detection of single biomarkers new proteomic-based high-throughput profiling techniques were developed. Proteome analysis of body fluids, including serum, ascites and pleural effusions (10) can provide the link between gene sequence and cellular physiology (46). Blood serum is the main subject of biomarker research due to its accessibility. It also allows a dynamic and accurate reflection of the intrinsic genetic status of a cell and its environment (47). Initial proteomic approaches in ovarian cancer biomarker research have used SELDI-mass spectrometry or immobilized metal affinity chromatography (IMAC)-nickel chip surface to investigate a sub-proteome of metal-binding proteins for detection of biomarkers, and has shown better sensitivity and specificity compared to CA125 (48-50). No individual candidate protein which were identified in these studies has shown better sensitivity and specificity compared to CA125. The bioinformatically combined use of selected biomarkers and CA125 yielded a comparable or slightly higher sensitivity and specificity than CA125 alone. Nevertheless, to our knowledge no single biomarker identified by proteome-based profiling has shown better sensitivity and specificity compared to CA125.

Despite a broad spectrum of technologies, identified single biomarkers seem to reach a sensitivity and specificity plateau which is similar to CA125. Multivariate logistic regression analysis on various combinations of markers aimed to overtake the performance of CA125, and various studies have used a panel of markers with or without CA125. Multivariate combination of three biomarkers combined with CA125 gained a higher sensitivity and specificity than CA125 alone (51). Classification regression trees have been developed, and a combination of five markers (CA125, OVX1, LASA, CA15-3, CA72-4) detected by 'Classification and regression tree analysis' reached a high sensitivity (90.6%) and specificity (93.2%) (52), whilst discrimination of malignant from benign pelvic mass by artificial neural network analysis (ANN) revealed an even increased test specificity due to only four (CA125, LASA, CA15-3,

CA72-4) of these markers, reaching 20% higher specificity and similar sensitivity as CA125 alone (53).

In our biomarker discovery approach using printed glycan array, CA125 showed comparable sensitivity and specificity with the detection of a naturally occurring AGA recognizing P_1 glycan. We also used a multivariate selection model to search for potential glycans that could be combined to improve our differentiation rate. The highest ranked glycans were therefore further combined in a linear discriminant model and the performance was investigated by ROC analysis. Panels of AGA did not, however, reach a better sensitivity and specificity than 84%. When using a combination of eight candidates containing the structure *Galβ1-4GlcNAcβ* in the top ten selected glycans, a sensitivity of 83.3% and specificity of 87.8% could be reached, which is clearly better than CA125 (76%/ 73.9%, respectively) in the same cohort. As our results are generated from one single Swiss cohort, we will next validate these findings in a second independent Australian cohort of over 350 patients with gynaecological adenocarcinoma.

Other more clinically-based models like the 'risk of malignancy index' (RMI) (54) were using menopausal status, ultrasound and CA125 measurements in combination. The RMI is the product of CA125 measurement, ultrasound scan results (expressed as a score of 0, 1 or 3) and the menopausal status (1 if premenopausal and 3 if postmenopausal). RMI has been shown to be able to distinguish benign and malignant cases better than CA125 alone (54). Another study investigated preoperative serum levels of HE4 and CA125, combining them in a logistic regression algorithm with pre- and postmenopausal status (25). This algorithm classified patients into low and high risk groups with an accuracy of 93.8% (25). Although these clinical tests can be used to triage patients for gynaecological oncology centres in the presence of a palpable pelvic mass, there is no benefit given from screening asymptomatic women with either ultrasound or CA125.

4.2 **Biological and Technical Complexity of Saccharides**

4.2.1 High-Throughput Glycan Arrays

One of the current frontiers in "omic" sciences these days is the detection of the glycome, namely carbohydrate structures which are linked to proteins and lipids within an individual organism. High-throughput profiling of glycan-binding biomolecules by carbohydrate based microarrays allows determination of specificities of glycan-binding proteins (GBP) (55), examination of microbiologically relevant glycans (56), and the study of

carbohydrate-processing enzymes (57) and anti-glycan immune response (58, 59). Microarray-based analysis of GBP in complex biological fluids are complicated but of highest biological and clinical relevance. To date, only few publications investigate a small cohort of glycans or glycoconjugates within a specific biological system (59), screening of AGA within human serum for diagnostic purposes or detecting the immune response to bacterial pathogens (58). The printed glycan array platform used in our discovery approach has been fabricated according to the technique of the Consortium for Functional Glycomics (www.functionalglycomics.org), and comprises about 200 mono- and oligosaccharides including the observed terminal oligosaccharides of mammalian *N*- and *O*-glycoprotein chains, glycolipids, and their related cores. About twenty of these oligosaccharides are thought to be tumor-associated (60), and therefore thought to be synthesized by cancer cells in a more intense way than by normal cells.

4.2.2 Complexity of Glycan Array Fabrication

Various biological systems consist of a high diversity of glycans from monosaccharides up to 500'000-mer saccharide structures (61). Glycosylation processes include more than 50 glycosyltransferases, glycosidases and sugar nucleotide transporters (62). With only ten monosaccharides as a basis, $15x10^6$ tetramers (63) and $1.05x10^{12}$ hexamers can be built, thus demonstrating the complexity of the glycome (64). This enormous natural complexity is currently investigated by new technologies which are able to profile glycans and their binding partners.

Two major obstacles have hindered the recent development of glycan arrays, namely the difficult process of carbohydrate synthesis and the technique of immobilizing these structures onto an appropriate surface. A library of glycans for use in a high-throughput glycan array can be created by either carbohydrate synthesis or isolation of glycans from natural sources. Synthesis of carbohydrates can be achieved hereby with either automated solid-phase synthesis (65-67) or programmable "one spot" approach (68, 69). In case of enzymatic synthesis, glycosyltransferases and glycosidases are used (70), however, specific glycan processing enzymes are restricted in their availability due to their limited identification, characterization and their adaptability to synthesize carbohydrates enzymatically. The printed glycan array which we used for our discovery approach was based on monosaccharide and complex oligosaccharide structures synthesized by unique bioorganic pathways, thus achieving a final oligosaccharide purity of approximately 98% (71-73) (Lectinity Holdings, Moscow). Each glycan has been synthesized by a special and novel multi-step procedure, and was

confirmed by thin layer chromatography with the final product being characterized by nuclear magnetic resonance spectroscopy and mass spectrometry.

Glycan-based microarrays utilize different ways of glycan immobilization such as NHS-activated glass slides or nitrocellulose. Carbohydrate structures can be linked to different platforms by various chemical bonds and spacers and can be coupled covalently or non-covalently, specific or non-specific, and directly to the chemical surface or indirectly *via* glycoconjugates (74). As a type of planar surface, type NHS-activated hydrogel-coated glass slides were used in the printed glycan array to bind covalently amine-modified saccharides (75, 76). Using covalent printing, all monomeric glycans were attached in known and identical orientations; however, in contrast to polymeric presentation this arrangement limits probably anti-glycan antibody binding sites.

4.2.3 Problems of Standards and Controls

Tumor marker CA125 has been established for the detection of EOC for more than 20 years. It is a heavily glycosylated mucin identified by a murine mAb that lacks standardization based on poor purification, which is why all CA125 immunoassay results are expressed in U/mL. To increase performance of CA125 detection, a sandwich ELISA using two mAb was developed (77). An epitope-mapping workshop conducted by the International Society of Onco-Developmental Biology and Medicine investigated the binding characteristics of mAb reactive with CA125 and found that the binding affinity can be segregated into two major regions, OC125-like and M11-like epitopes (78). CA125 antigen is represented by a heterogeneous mixture of glycoproteins ranged in their molecular weight from 200 to 1000kDa, probably related to differences in glycosylations. It is expected that mouse and rat mAb preferentially bind to peptide epitopes located at the C-terminus of the protein. Based on non-existing standards and controls it is difficult to study glycan-protein interactions in this setting, which is a well known problem in glycobiology.

We used α-rhamnose as the biological positive control in our discovery approach. High expression levels of anti- α-rhamnose antibodies were observed in an IgG affinity-purified pool of 10'000 healthy individuals (79). Moderate to high levels of anti- α-rhamnose antibodies of immunoglobulin subtypes IgG, IgA and IgM were found in 200 individual healthy serum samples (79), consistent with our results. Fluorescence signals regarding α-rhamnose are within the 30% of highest glycan signals. The use of α-rhamnose as a positive biological control is further supported by findings that related antibodies have moderate biological variability of approximately 50% and a low coefficient of variation (80, 81).

Aminoglucitol has been suggested as negative biological control. It is an opened reduced form of *D*-glucose which is not present as a structural component of regular glycosylation. It has also been negative for anti-glycan binding at similar values to the technical background binding control (82), and was found in our study within the 18% lowest glycans. Glycan array technology has not established yet any standardized controls and the use of the described positive and negative biological controls are therefore only the best possible controls as to the literature.

4.2.4 Reproducibility of Glycan Arrays

Reproducibility of glycan array data was so far not studied except in one single publication where experimental variability, inter-individual biological variability and intra-individual temporal variability was investigated (80). An overall coefficient of variation (CV) of 10.8% and variation across multiple batches of slides and days of 28.5% was found, which is consistent with our results, despite different platforms and presentation of glycans to human antibodies. We also examined intra-slide variability using the glycan *Fucα1-2(GalNAcβ1-3)Galβ*, presented twice in two different positions on the array platform. A CV of less than 10% was observed in each experimental set up, demonstrating excellent reliability of results. Despite indications of a good adaptability of glycan arrays, further efforts must be undertaken to study human AGA. This includes further studies on clinical parameters into dependency of AGA patterns, studying antibody glycan-binding and data processing.

4.3 **Naturally Occurring Anti-Glycan Antibodies**

4.3.1 Studying Glycan Binding Motifs

Most glycan-binding proteins are able to recognize epitopes including two or three monosaccharides (83), and the non-glycan moiety of glycoconjugates seem to be involved in this binding. The study of GBP is a major field of glycobiology which is necessary to understand molecular recognition events. Measurements of glycan binding specificities is usually performed by molecular biological methods like affinity chromatography (62, 84, 85) and carbohydrate arrays (61, 74). One obvious disadvantage of affinity chromatography over glycan microarrays is the use of high amounts of glycans, including the enormous amount of time taken to synthesize carbohydrates to probe numerous glycan-protein interactions. High-throughput technologies allow profiling of GBP to a broad spectrum of glycans but are still challenging from the point of data interpretation and determination of specific complex carbohydrate binding motifs.

So far, only a few publications have studied glycan-binding motifs. Only one systematic analysis was performed to study the complexity of oligosaccharides, multiple specificities of GBP and the nature of glycan array data (86). Hereby, glycan motifs that are critical for binding were studied by bioinformatics using data sets produced by the Consortium for Functional Glycomics (86). First, lectin binding to each glycan was analyzed in terms of the motifs in order to identify motifs that are selectively present in glycans bound by the lectin. Two different methods to calculate the motifs were compared and validated using three different lectins; *Vicia villosa* lectin recognizes terminal *GalNAc*, Concanavalin A recognizes terminal mannose containing structures, and wheat germ agglutinin which usually binds to a broad and not yet fully understood spectrum of oligosaccharides. Both the 'motif segregation method' and the 'intensity segregation' identified accurately the primary binding specificities and secondary specificities of all three lectins. The feasibility of these systematic methods led to further investigations on motif specificities of plant and human lectins and the results revealed unexpected carbohydrate ligands (86, 87).

Only a few attempts were made to investigate the binding motifs of naturally occurring AGA. Motif specificities of AGA revealed higher cross-reactivity and a broader spectrum of glycan-binding motifs (86). Another study confirmed these finding by screening 27 well-defined AGA bound to 80 different glycans and glycoconjugates on a microarray platform. More than a half of the profiled monoclonal AGA bound not only with designated antigens but demonstrated a high degree of cross reactivity. (88). In our experiments human polyclonal AGA seemed to preferentially bind to specific core structures. We observed that binding affinity or motif specificity of AGA seemed higher in the case of such core structures, for example *GalNAcα-*, *Galβ1-4GlcNAc-* or *Galβ1-4Glcβ-*. In contrast, no binding motifs could be observed on monosaccharide components possibly related to an overlay of unspecific binding.

Unfortunately, a pre-processing of glycan microarray data into biological motifs has been performed and thus does not allow the comparison of data to our results (86). Glycan-recognition epitopes show anti-Le^b- antibodies (25LE) binding to Le^b, Le^a and terminal *Fucα1-4* (86). Strong signals can be also found for anti-Le^b antibodies to core structures, terminal *GlcNAcβ* as well as *GlcNAcβ1-6Gal* which is not primarily presented at the terminal part of Le^b oligosaccharide (88). Moreover, high cross-reactivities of AGA observed in both studies (86, 88) are in concordance to our observations.

Specific AGA binding was used to monitor the expression of carbohydrate antigens, primarily on tissue sections to modulate their biological activity, and to target specific carbohydrates for clinical applications. Conclusions based on these findings are frequently

based on the perception that AGA are highly specific for their listed glycan antigens. Using glycan arrays we have shown that naturally occurring AGA of different immunoglobulin subtypes do not have such glycan-binding specificities. We are the first suggesting that naturally occurring AGA seem to bind epitope structures mostly located at the core of a complex carbohydrate chain.

4.3.2 Anti-Glycan Antibodies Bind to Specific Carbohydrate Core Structures

Various functional studies have examined naturally occurring AGA and their binding patterns to glycans, namely anti-α-galactosyl antibodies (89, 90), antibodies bound to negatively charged glycans (82, 91), and *GlcNAc*-terminated carbohydrate chains (82, 92). Carbohydrates are exposed on the cell surface and on secreted proteins, acting as sensors for biochemical signals (93). In a previous study, human naturally occurring AGA recognized short inner core structures for glycolipids (*Galβ1-4Glc*) and glycoproteins (*GalNAcα*) (82). The core structure is defined as a carbohydrate fragment located at the onset of a complex oligosaccharide, namely, the first carbohydrate structure at the linkage to a related glycolipid or –protein. We identified a panel of AGA core structures which occur in specific patterns and could describe 21 key core carbohydrate structures consisting of a monosaccharide (*e.g. GalNAcα*) up to a tetrasaccharide (*e.g. Galβ1-3GlcNAcβ1-3Galβ1-4Glcβ*). Terminally located monosaccharides (*e.g. Galα1-3(Fucα1-2)Galβ*) or terminal modifications (*e.g. 6'-O-Su-Galβ1-4GlcNAc*) seem to be able to influence binding of AGA to core structures. These monosaccharides are important for cell-cell interactions, as has been shown in host-pathogen recognition (94) and terminally located *αGal* (89, 90). Interestingly, comparison of healthy controls and non-mucinous EOC serum samples by multivariate analysis revealed a panel of glycans with similar core patterns responsible for significant discrimination between both subgroups. By analysing the first ten 'top candidate' structures, 8 out of 10 oligosaccharides contained the inner core structure -*Galβ1-4GlcNAc* and 2 oligosaccharides contained -*GalNAcα*. These observations led us to propose that the majority of human AGA immunoglobulin subtypes IgA, IgG and IgM recognize inner core structures of cell surface glycosylations including *N*-, *O*-linked glycosylation and glycolipids. In non-malignant conditions inner core structures of glycosylation trees are occupied by interactions of different types of membrane-associated proteins and other carbohydrate structures (Figure 23 A). At this stage, inner core structures are encrypted and the binding of naturally occurring AGA seem to be limited to these potential cell surface carbohydrate core structures. In contrast, oncogenic transformation results in rearrangement of extra-cellular matrix and their associated molecular

components. Thus, encrypted binding sites could become available for circulating AGA, thus decrypting inner core structures (Figure 23 B). Alteration in glycan presentation signalizes the status of the cell towards the innate immune system; mainly to B1 cells which produce naturally occurring AGA.

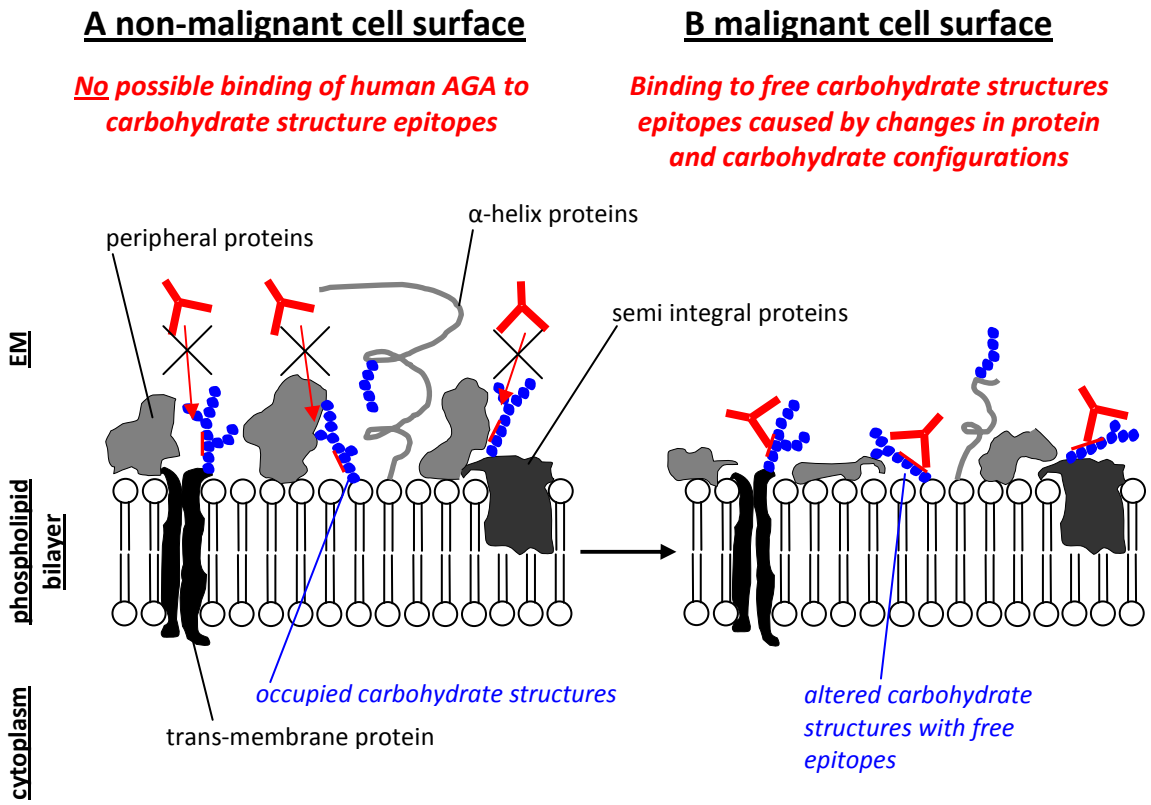


Figure 23 Anti-glycan antibody binding to inner core structures of membrane associated oligosaccharide chains.
In non-malignant cell surface membranes (A) binding of AGA to core epitopes (red line) is not possible due to occupation of surrounding proteins (grey, black) - inner core structures are encrypted. During oncogenic transformation cell surface composition changes (B) and core carbohydrate structures are recognizable by circulating AGA- inner core structures are decrypted. EM= Extracellular Matrix

4.3.3 Influence of Monomeric *versus* Polymeric Presented Glycans

The presentation of glycans is of major importance for the investigation of naturally occurring AGA binding specificities. Most studies investigated glycan-binding patterns using one examination platform, which in our opinion bears danger for conclusions. In one of our studies we applied three independent custom-developed 'glycan-based immunoassays' to detect human AGA for methodological and statistical comparison of assays and to be able to interpret the biology based on AGA detection. We compared PGA, a high-throughput approach (76, 82), to multiplex flow cytometric suspension array (SA) and standard ELISA. PGA has the advantage of high sensitivity, low background signals, screening of an unlimited repertoire of

natural and synthesized glycans, speed and small reagent consumption. Multiplex flow cytometric SA (95) incorporates fluorescent microspheres with distinct spectral addresses as a solid support for glycans (58). This type of assay combines the flexibility of routine ELISA with simultaneous detection of multiple ligands in one sample and minimal reagent consumption. ELISA is widely used in most immunochemical studies (96-99), is affordable and well established method for screening of limited number of ligands.

Binding of antibody subtypes IgA, IgG and IgM to specific glycans was measured in all three immunoassays, and signals detected using different detection methods; streptavidin HRPO and TMB (ELISA), Streptavidin Alex Fluor^{555} (PGA), and R-phycoerythrin labeled secondary antibodies (SA). Although all techniques aim to measure the same chemically synthesized glycans presented to AGA, they are captured differently, monomeric on flat surface/ 2-dimensional (PGA), polymeric on microspheres/ 3-dimensional (SA), and polymeric on flat surface/ semi 2-dimensional (ELISA). In our comparative analysis, distribution of AGA to blood group related trisaccharides showed comparable binding parameters. The pattern of AGA binding changed when screening for the PGA derived ovarian cancer biomarker trisaccharide P_1 which was measured in all three techniques. The marked discrepancy in detecting P_1 in all methods could be caused by various factors: a) potential cancer association of P_1 antibody levels as observed in our previous study, b) low number of serum samples in study cohort, c) differences in assay conditions like glycan concentration and serum dilution. Concordance correlation coefficient among assays was much stronger in the case of AB0 blood group antigens than in the case of P_1. This could be explained by the strong biological association and high specificity of anti- A/B antibodies (100). Affinity and biological significance of anti-P_1 antibodies has not yet been clearly investigated, however, they obviously represent a subpopulation of AGA with lower affinity compared to anti-A/B antibodies. In the case of PGA, glycans are presented as monomers in a highly organized manner with limited flexibility; only a subpopulation of anti-P_1 specific antibodies could be possibly detected. Thus, it is necessary to highlight that in contrast to PGA, SA and ELISA are assays based on polymeric glycan presentation which means that multivalent binding between glycans and high-affinity AGA can here possibly reflect a broader spectrum of glycan antibody interactions (58, 95). Non-specific binding such as direct binding of heterophilic serum antibodies in SA and ELISA could interfere with specific signals, as previously stated (101, 102). We proposed that the assays based on multivalent and polymeric ligand presentation seem to be more close to natural or *in vivo* conditions. For biomarker research, in order to identify markers of disease, the advantage of using one assay with appropriate sensitivity,

specificity and reproducibility might be sufficient. For a rather functional approach into the biology of detected AGA, each glycan detected should be investigated further using also other immunoassays. Glycan-binding interactions can also be studied nuclear magnetic resonance spectroscopy of affinity-purified antibodies or by advanced surface plasmon resonance technology as a real time measurement.

4.4 The Impact of Carbohydrates in Biomedical Research

Glycans have been underestimated by the scientific community for a long time, whilst research focused on nucleotides, DNA, peptides and proteins. This can be explained by the complexity of carbohydrate synthesis and the fact that glycans are products of non-template driven processes because they are products of a network of glycosylation-related enzymes. With an increasing interest in glycobiology during the last couple of years an increasing number of publications in the field of organic chemistry, biochemistry, molecular biology, and biomedicine have arisen. As glycans play a major role in many biological processes, a large interest has driven the development of (1) new technologies to study carbohydrate-biomolecule interactions, (2) new potential diagnostic and therapeutic agents, (3) eventually clinical applications (Figure 24).

Glycans are ubiquitous and represent the most abundant class of molecules in the nature. They coat all cell surfaces in a broad complexity to act as recognition molecules and therefore are involved in pathological phenotypes such as inflammation, cancer, and infectious diseases. There is a recognized need for glycobiology-driven therapeutic applications, thus the investigation of functional carbohydrate epitopes and its protein receptor (*e.g.* lectin, AGA) is eminent. Technologies allowing such investigations include nuclear magnetic resonance spectroscopy which determines the exact binding conformation and X-ray crystallography of the carbohydrate ligand, co-crystallized to its GBP. Synthesized glyco-mimicry structures of currently approved drugs are glycosidase inhibitors that prevent the digestion of carbohydrates for the prevention of influenza virus infections (Relenza, Tamiflu), the treatment of diabetes (Glustat, Zavesca, Glyset, Glucobay), and sulphated glycosaminoglycans which function as anticoagulants by binding to antithrombin III for the treatment of thrombosis (Arixtra) (103). Successful conjugations of carbohydrates to nanoparticles have attracted significant interest in recent years due to their specific properties. In addition to the unique properties originating from their quantum structure, nanoparticles display a large surface area and thus provide an ideal platform that can be conjugated with biomolecules for targeted approaches for cancer diagnostics, therapeutics and imaging (104). Carbohydrate-functionalized magnetic

nanoparticles were also used to detect the endothelial markers E-/ P-selectins, which are associated with acute inflammation of the cerebral endothelium (105). Imaging of metabolic carbohydrates to localize or distinguish the target from the surrounding components by using spectroscopic probes is becoming increasingly important. Whilst currently genetically encoded tags such as GFP are routinely used for protein trafficking, studies have demonstrated the use of metabolically labelled cell-surface sialic acids (106).

The role of carbohydrate structures conjugated to cancer vaccines has been studied intensively. An overexpression of ganglioside GD3 has shown to be associated to melanomas, but gangliosides were of poor immunogenicity. Therefore, GD3 has been conjugated to an immunogenic carrier, keyhole limpet hemocyanin (KLH), demonstrating stronger antibody immune responses in immunized mice (107). For enhancement of the immune response, applications of adjuvants can significantly increase the immune response of patient to a co-administered antigen (108). A purified plant extract derived from soap bark tree *Quillaja saponaria* (QS-21) was the most effective adjuvant and shows good adaptability to the IgM and IgG antibody response against MUC1 and GD3 (109), currently being under clinical evaluation as adjuvant for trial vaccines including HIV, malaria and cancer.

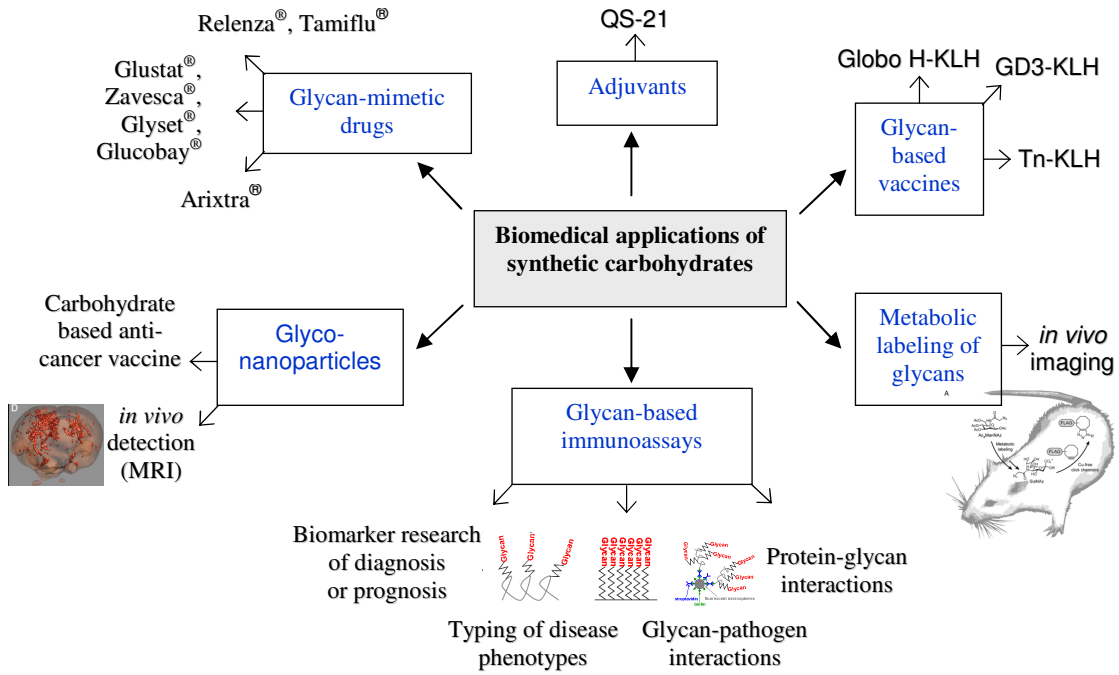


Figure 24 Overview of biomedical applications of synthetic carbohydrates.
Synthetic carbohydrates are multi-faceted tools; glycans were applied as glycan-mimetic drugs, coupled to nanoparticles for *in vivo* imaging of GBP (108), and are useful as adjuvants to enhance human immune response. Glycan-based vaccines (113) as well as metabolic labeling of glycans has recently been documented (109).

4.5 References

1. Ozols RF (2006) Challenges for chemotherapy in ovarian cancer. *Ann Oncol* 17 Suppl 5:v181-187.
2. Fishman DA & Bozorgi K (2002) The scientific basis of early detection of epithelial ovarian cancer: the National Ovarian Cancer Early Detection Program (NOCEDP). *Cancer Treat Res* 107:3-28.
3. Agarwal R, Linch M, & Kaye SB (2006) Novel therapeutic agents in ovarian cancer. *Eur J Surg Oncol* 32(8):875-886.
4. Carr KM, Rosenblatt K, Petricoin EF, & Liotta LA (2004) Genomic and proteomic approaches for studying human cancer: prospects for true patient-tailored therapy. *Hum Genomics* 1(2):134-140.
5. Afshari CA, Nuwaysir EF, & Barrett JC (1999) Application of complementary DNA microarray technology to carcinogen identification, toxicology, and drug safety evaluation. *Cancer Res* 59(19):4759-4760.
6. Duan Z, Feller AJ, Penson RT, Chabner BA, & Seiden MV (1999) Discovery of differentially expressed genes associated with paclitaxel resistance using cDNA array technology: analysis of interleukin (IL) 6, IL-8, and monocyte chemotactic protein 1 in the paclitaxel-resistant phenotype. *Clin Cancer Res* 5(11):3445-3453.
7. Schummer M*, et al.* (1999) Comparative hybridization of an array of 21,500 ovarian cDNAs for the discovery of genes overexpressed in ovarian carcinomas. *Gene* 238(2):375-385.
8. Wang K*, et al.* (1999) Monitoring gene expression profile changes in ovarian carcinomas using cDNA microarray. *Gene* 229(1-2):101-108.
9. Naora H, Montz FJ, Chai CY, & Roden RB (2001) Aberrant expression of homeobox gene HOXA7 is associated with mullerian-like differentiation of epithelial ovarian tumors and the generation of a specific autologous antibody response. *Proc Natl Acad Sci U S A* 98(26):15209-15214.
10. Jacob F, Goldstein DR, Fink D, & Heinzelmann-Schwarz V (2009) Proteogenomic studies in epithelial ovarian cancer: established knowledge and future needs. *Biomarkers Med.* 3(6):743-756.
11. Mocellin S, Wang E, Panelli M, Pilati P, & Marincola FM (2004) DNA array-based gene profiling in tumor immunology. *Clin Cancer Res* 10(14):4597-4606.
12. Barrett T*, et al.* (2005) NCBI GEO: mining millions of expression profiles--database and tools. *Nucleic Acids Res* 33(Database issue):D562-566.
13. Parkinson H*, et al.* (2009) ArrayExpress update--from an archive of functional genomics experiments to the atlas of gene expression. *Nucleic Acids Res* 37(Database issue):D868-872.
14. Larsson O & Sandberg R (2006) Lack of correct data format and comparability limits future integrative microarray research. *Nat Biotechnol* 24(11):1322-1323.
15. Brazma A*, et al.* (2001) Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nat Genet* 29(4):365-371.
16. Zimmermann P, Hirsch-Hoffmann M, Hennig L, & Gruissem W (2004) GENEVESTIGATOR. Arabidopsis microarray database and analysis toolbox. *Plant Physiol* 136(1):2621-2632.
17. Rhodes DR*, et al.* (2004) ONCOMINE: a cancer microarray database and integrated data-mining platform. *Neoplasia* 6(1):1-6.
18. Cochrane AL, Maynard A, & Chalmers I (1997) *Non-random reflections on health services research: on the 25th anniversary of Archie Cochrane's Effectiveness and efficiency* (BMJ, London) pp xv, 303.

19. Lancaster JM, *et al.* (2004) Gene expression patterns that characterize advanced stage serous ovarian cancers. *J Soc Gynecol Investig* 11(1):51-59.
20. Schaner ME, *et al.* (2003) Gene expression patterns in ovarian carcinomas. *Mol Biol Cell* 14(11):4376-4386.
21. Schwartz DR, *et al.* (2002) Gene expression in ovarian cancer reflects both morphology and biological behavior, distinguishing clear cell from other poor-prognosis ovarian carcinomas. *Cancer Res* 62(16):4722-4729.
22. Welsh JB, *et al.* (2001) Analysis of gene expression profiles in normal and neoplastic ovarian tissue samples identifies candidate molecular markers of epithelial ovarian cancer. *PNAS* 98:1176-1181.
23. Dressman HK, *et al.* (2007) An integrated genomic-based approach to individualized treatment of patients with advanced-stage ovarian cancer. *J Clin Oncol* 25(5):517-525.
24. Galgano MT, Hampton GM, & Frierson HF (2006) Comprehensive analysis of HE4 expression in normal and malignant human tissues. *Mod Pathol* 19(6):847-853.
25. Moore RG, *et al.* (2009) A novel multiple marker bioassay utilizing HE4 and CA125 for the prediction of ovarian cancer in patients with a pelvic mass. *Gynecol Oncol* 112(1):40-46.
26. Hellstrom I, *et al.* (2003) The HE4 (WFDC2) protein is a biomarker for ovarian carcinoma. *Cancer Res* 63(13):3695-3700.
27. Heinzelmann-Schwarz VA, *et al.* (2004) Overexpression of the cell adhesion molecules DDR1, Claudin 3, and Ep-CAM in metaplastic ovarian epithelium and ovarian cancer. *Clin Cancer Res* 10(13):4427-4436.
28. Taylor CF, *et al.* (2007) The minimum information about a proteomics experiment (MIAPE). *Nat Biotechnol* 25(8):887-893.
29. Deutsch EW, *et al.* (2008) Minimum information specification for in situ hybridization and immunohistochemistry experiments (MISFISHIE). *Nat Biotechnol* 26(3):305-312.
30. Deutsch EW, *et al.* (2006) Development of the Minimum Information Specification for In Situ Hybridization and Immunohistochemistry Experiments (MISFISHIE). *Omics* 10(2):205-208.
31. Taylor CF, *et al.* (2008) Promoting coherent minimum reporting guidelines for biological and biomedical investigations: the MIBBI project. *Nat Biotechnol* 26(8):889-896.
32. Timmons JA, *et al.* (2005) Human muscle gene expression responses to endurance training provide a novel perspective on Duchenne muscular dystrophy. *Faseb J* 19(7):750-760.
33. Zhou XJ, *et al.* (2005) Functional annotation and network reconstruction through cross-platform integration of microarray data. *Nat Biotechnol* 23(2):238-243.
34. Fehrmann RS, *et al.* (2007) Profiling studies in ovarian cancer: a review. *Oncologist* 12(8):960-966.
35. Benowitz S (2008) Rethinking cancer vaccine trials: would new measures of success make a difference? *J Natl Cancer Inst* 100(4):237-238.
36. Ashworth A, *et al.* (2008) Opportunities and challenges in ovarian cancer research, a perspective from the 11th Ovarian cancer action/HHMT Forum, Lake Como, March 2007. *Gynecol Oncol* 108(3):652-657.
37. Turnbull JE & Field RA (2007) Emerging glycomics technologies. *Nat Chem Biol* 3(2):74-77.
38. Bingle L, Singleton V, & Bingle CD (2002) The putative ovarian tumour marker gene HE4 (WFDC2), is expressed in normal tissues and undergoes complex altrnative splicing to yield multiple protein isoforms. *Oncogene* 21:2768-2773.

39. Hellstrom I & Hellstrom KE (2008) SMRP and HE4 as biomarkers for ovarian carcinoma when used alone and in combination with CA125 and/or each other. *Adv Exp Med Biol* 622:15-21.
40. Bast RC, Jr.*, et al.* (2005) New tumor markers: CA125 and beyond. *Int J Gynecol Cancer* 15 Suppl 3:274-281.
41. Kim JH*, et al.* (2002) Osteopontin as a potential diagnostic biomarker for ovarian cancer. *Jama* 287(13):1671-1679.
42. Carpini JD, Karam AK, & Montgomery L (Vascular endothelial growth factor and its relationship to the prognosis and treatment of breast, ovarian, and cervical cancer. *Angiogenesis* 13(1):43-58.
43. Sunde JS*, et al.* (2006) Expression Profiling Identifies Altered Expression of Genes That Contribute to the Inhibition of Transforming Growth Factor-{beta} Signaling in Ovarian Cancer. *Cancer Res* 66(17):8404-8412.
44. Dong Y*, et al.* (1997) Expression of MUC1 and MUC2 mucins in epithelial ovarian tumours. *J Pathol* 183:311-317.
45. Raki M, Rein DT, Kanerva A, & Hemminki A (2006) Gene transfer approaches for gynecological diseases. *Mol Ther* 14(2):154-163.
46. Dove A (1999) Proteomics: translating genomics into products? *Nat Biotechnol* 17(3):233-236.
47. Aebersold R*, et al.* (2005) Perspective: a program to improve protein biomarker discovery for cancer. *J Proteome Res* 4(4):1104-1109.
48. Petricoin EF*, et al.* (2002) Use of proteomic patterns in serum to identify ovarian cancer. *Lancet* 359(9306):572-577.
49. Rai AJ*, et al.* (2002) Proteomic approaches to tumor marker discovery. *Arch Pathol Lab Med* 126(12):1518-1526.
50. Ye B*, et al.* (2003) Haptoglobin-alpha subunit as potential serum biomarker in ovarian cancer: identification and characterization using proteomic profiling and mass spectrometry. *Clin Cancer Res* 9(8):2904-2911.
51. Zhang Z*, et al.* (2004) Three biomarkers identified from serum proteomic analysis for the detection of early stage ovarian cancer. *Cancer Res* 64(16):5882-5890.
52. Woolas RP, Oram DH, Jeyarajah AR, Bast RC, & Jacobs IJ (1999) Ovarian cancer identified through screening with serum markers but not by pelvic imaging. *Int J Gynecol Cancer* 9(6):497-501.
53. Zhang Z*, et al.* (1999) Combination of multiple serum markers using an artificial neural network to improve specificity in discriminating malignant from benign pelvic masses. *Gynecol Oncol* 73(1):56-61.
54. Jacobs I*, et al.* (1990) A risk of malignancy index incorporating CA 125, ultrasound and menopausal status for the accurate preoperative diagnosis of ovarian cancer. *Br J Obstet Gynaecol* 97(10):922-929.
55. Moller I*, et al.* (2008) High-throughput screening of monoclonal antibodies against plant cell wall glycans by hierarchical clustering of their carbohydrate microarray binding profiles. *Glycoconj J* 25(1):37-48.
56. Disney MD & Seeberger PH (2004) The use of carbohydrate microarrays to study carbohydrate-cell interactions and to detect pathogens. *Chem Biol* 11(12):1701-1707.
57. Park S & Shin I (2007) Carbohydrate microarrays for assaying galactosyltransferase activity. *Org Lett* 9(9):1675-1678.
58. Bovin NV & Huflejt ME (2008) Unlimited glycochip. *Trends Glycosci. Glycotechnol.* 20(115):245-258.
59. Krishnamoorthy L & Mahal LK (2009) Glycomic analysis: an array of technologies. *ACS Chem Biol* 4(9):715-732.

60. Hakomori S (2001) Tumor-associated carbohydrate antigens defining tumor malignancy: basis for development of anti-cancer vaccines. *Adv Exp Med Biol* 491:369-402.
61. Wang D (2003) Carbohydrate microarrays. *Proteomics* 3(11):2167-2175.
62. Hirabayashi J, Arata Y, & Kasai K (2003) Frontal affinity chromatography as a tool for elucidation of sugar recognition properties of lectins. *Methods Enzymol* 362:353-368.
63. Dove A (2001) The bittersweet promise of glycobiology. *Nat Biotechnol* 19(10):913-917.
64. Laine RA (1994) A calculation of all possible oligosaccharide isomers both branched and linear yields 1.05 x 10(12) structures for a reducing hexasaccharide: the Isomer Barrier to development of single-method saccharide sequencing or synthesis systems. *Glycobiology* 4(6):759-767.
65. Palmacci ER, Hewitt MC, & Seeberger PH (2001) "Cap-Tag"-Novel Methods for the Rapid Purification of Oligosaccharides Prepared by Automated Solid-Phase Synthesis Financial support from the donors of the Petroleum Research Fund, administered by the ACS (ACS-PRF 34649-G1), Merck (Predoctoral Fellowship for E.R.P.), Boehringer-Ingelheim (Predoctoral Fellowship for E.R.P.), and the NIH (Biotechnology Training Grant for M.C.H.) is gratefully acknowledged. Funding for the MIT-DCIF Inova 501 was provided by the NSF (Award CHE-9808061). Funding for the MIT-DCIF Avance (DPX) 400 was provided by the NIH (Award 1S10RR13886-01). We thank Silicycle (Quebec City, Canada) for the generous gift of fluorous silica gel and isocyanate silica gel scavenger resin. We thank Dr. O. Plante for helpful discussions and help with the automated oligosaccharide synthesizer. *Angew Chem Int Ed Engl* 40(23):4433-4437.
66. Plante OJ, Palmacci ER, & Seeberger PH (2001) Automated solid-phase synthesis of oligosaccharides. *Science* 291(5508):1523-1527.
67. Seeberger PH & Haase WC (2000) Solid-phase oligosaccharide synthesis and combinatorial carbohydrate libraries. *Chem Rev* 100(12):4349-4394.
68. Sears P & Wong CH (2001) Toward automated synthesis of oligosaccharides and glycoproteins. *Science* 291(5512):2344-2350.
69. Zhang ZG, Lu JZ, & Chen JW (1999) Concanavalin A binding to oligosaccharide chain leads to alterations in properties of band 3. *Biochem Mol Biol Int* 47(3):377-385.
70. MacKenzie DA*, et al.* (1998) Aberrant processing of wild-type and mutant bovine pancreatic trypsin inhibitor secreted by Aspergillus niger. *J Biotechnol* 63(2):137-146.
71. Pazynina G*, et al.* (2009) Chemical synthesis of 6(GlcNAc)- and 6(Gal)-O-sulfated SiaLe(X) tetrasaccharides in spacer-armed form. *Glycobiology* 19(10):1078-1081.
72. Severov VV, Belianchikov IM, Pazynina GV, & Bovin NV (2007) [Synthesis of N-acetyllactosamine-containing oligosaccharides, galectin ligands]. *Bioorg Khim* 33(1):131-147.
73. Zemlianukhina TV & Bovin NV (1990) [Synthesis of oligosaccharides with group specificity for blood types A, B and H (Type 3)]. *Bioorg Khim* 16(8):1096-1104.
74. Culf AS, Cuperlovic-Culf M, & Ouellette RJ (2006) Carbohydrate microarrays: survey of fabrication techniques. *Omics* 10(3):289-310.
75. Shilova NV*, et al.* (2008) 2-Aminopyridine--a label for bridging of oligosaccharides HPLC profiling and glycoarray printing. *Glycoconj J* 25(1):11-14.
76. Blixt O*, et al.* (2004) Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. *Proc Natl Acad Sci U S A* 101(49):17033-17038.
77. Bast RC, Jr.*, et al.* (1998) CA 125: the past and the future. *Int J Biol Markers* 13(4):179-187.

78. Nustad K, *et al.* (1996) Specificity and affinity of 26 monoclonal antibodies against the CA 125 antigen: first report from the ISOBM TD-1 workshop. International Society for Oncodevelopmental Biology and Medicine. *Tumour Biol* 17(4):196-219.
79. Dotan N, Altstock RT, Schwarz M, & Dukler A (2006) Anti-glycan antibodies as biomarkers for diagnosis and prognosis. *Lupus* 15(7):442-450.
80. Oyelaran O, McShane LM, Dodd L, & Gildersleeve JC (2009) Profiling Human Serum Antibodies with a Carbohydrate Antigen Microarray. *J Proteome Res*.
81. Schwarz M, *et al.* (2006) Serum anti-Glc(alpha1,4)Glc(alpha) antibodies as a biomarker for relapsing-remitting multiple sclerosis. *J Neurol Sci* 244(1-2):59-68.
82. Huflejt ME, *et al.* (2009) Anti-carbohydrate antibodies of normal sera: findings, surprises and challenges. *Mol Immunol* 46(15):3037-3049.
83. Varki A, *et al.* (1999) *Essentials of Glycobiology* (Cold Spring Harbor Labartory Press, Plainview (NY)) p p.41.
84. Hirabayashi J (2004) Lectin-based structural glycomics: glycoproteomics and glycan profiling. *Glycoconj J* 21(1-2):35-40.
85. Tateno H, Nakamura-Tsuruta S, & Hirabayashi J (2007) Frontal affinity chromatography: sugar-protein interactions. *Nat Protoc* 2(10):2529-2537.
86. Porter A, *et al.* (2010) A motif-based analysis of glycan array data to determine the specificities of glycan-binding proteins. *Glycobiology* 20(3):369-380.
87. Manimala JC, Roach TA, Li Z, & Gildersleeve JC (2006) High-throughput carbohydrate microarray analysis of 24 lectins. *Angew Chem Int Ed Engl* 45(22):3607-3610.
88. Manimala JC, Roach TA, Li Z, & Gildersleeve JC (2007) High-throughput carbohydrate microarray profiling of 27 antibodies demonstrates widespread specificity problems. *Glycobiology* 17(8):17C-23C.
89. Milland J, *et al.* (2007) Carbohydrate residues downstream of the terminal Galalpha(1,3)Gal epitope modulate the specificity of xenoreactive antibodies. *Immunol Cell Biol* 85(8):623-632.
90. Parker W, *et al.* (1999) Naturally occurring anti-alpha-galactosyl antibodies: relationship to xenoreactive anti-alpha-galactosyl antibodies. *Glycobiology* 9(9):865-873.
91. Lekakh IV, Bovin NV, Bezyaeva GP, & Poverenny AM (2001) Natural hidden autoantibodies react with negatively charged carbohydrates and xenoantigen Bdi. *Biochemistry (Mosc)* 66(2):163-167.
92. Briko NI, *et al.* (1997) Enzyme immunoassay kit for detecting antibodies to group-specific antigen of group A Streptococcus on the base of conjugated N-acetylglucosamine and its medical application. *Klin.La.Diagn.* 9:43-46.
93. Reuter G & Gabius HJ (1999) Eukaryotic glycosylation: whim of nature or multipurpose tool? *Cell Mol Life Sci* 55(3):368-422.
94. Chandrasekaran A, *et al.* (2008) Glycan topology determines human adaptation of avian H5N1 virus hemagglutinin. *Nat Biotechnol* 26(1):107-113.
95. Nolan JP & Sklar LA (2002) Suspension array technology: evolution of the flat-array paradigm. *Trends Biotechnol* 20(1):9-12.
96. Buchs JP & Nydegger UE (1989) Development of an ABO-ELISA for the quantitation of human blood group anti-A and anti-B IgM and IgG antibodies. *J Immunol Methods* 118(1):37-46.
97. Galanina OE, Mecklenburg M, Nifantiev NE, Pazynina GV, & Bovin NV (2003) GlycoChip: multiarray for the study of carbohydrate-binding proteins. *Lab Chip* 3(4):260-265.

98. Obukhova P, Rieben R, & Bovin N (2007) Normal human serum contains high levels of anti-Gal alpha 1-4GlcNAc antibodies. *Xenotransplantation* 14(6):627-635.
99. Rieben R, Buchs JP, Fluckiger E, & Nydegger UE (1991) Antibodies to histo-blood group substances A and B: agglutination titers, Ig class, and IgG subclasses in healthy persons of different age categories. *Transfusion* 31(7):607-615.
100. Watkins WM (2001) The ABO blood group system: historical background. *Transfus Med* 11(4):243-265.
101. Martins TB, Augustine NH, & Hill HR (2006) Development of a multiplexed fluorescent immunoassay for the quantitation of antibody responses to group A streptococci. *J Immunol Methods* 316(1-2):97-106.
102. Waterboer T, Sehr P, & Pawlita M (2006) Suppression of non-specific binding in serological Luminex assays. *J Immunol Methods* 309(1-2):200-204.
103. Magnani JL & Ernst B (2009) Glycomimetic drugs - a new source of therapeutic opportunities. *Discov Med* 8(43):247-252.
104. Nie S, Xing Y, Kim GJ, & Simons JW (2007) Nanotechnology applications in cancer. *Annu Rev Biomed Eng* 9:257-288.
105. van Kasteren SI*, et al.* (2009) Glyconanoparticles allow pre-symptomatic in vivo imaging of brain disease. *Proc Natl Acad Sci U S A* 106(1):18-23.
106. Chang PV*, et al.* (Copper-free click chemistry in living animals. *Proc Natl Acad Sci U S A* 107(5):1821-1826.
107. Helling F*, et al.* (1994) GD3 vaccines for melanoma: superior immunogenicity of keyhole limpet hemocyanin conjugate vaccines. *Cancer Res* 54(1):197-203.
108. Jiang ZH & Koganty RR (2003) Synthetic vaccines: the role of adjuvants in immune targeting. *Curr Med Chem* 10(15):1423-1439.
109. Kim SK*, et al.* (1999) Comparison of the effect of different immunological adjuvants on the antibody and T-cell response to immunization with MUC1-KLH and GD3-KLH conjugate cancer vaccines. *Vaccine* 18(7-8):597-603.
110. Slovin SF, Keding SJ, & Ragupathi G (2005) Carbohydrate vaccines as immunotherapy for cancer. *Immunol Cell Biol* 83(4):418-428.

5 Future Prospects

5.1 Cross-Validation of Printed Glycan Arrays

Tumor associated carbohydrate antigens (TACA) are known to trigger an immune response resulting in the generation of anti-glycan antibodies (AGA). Based on literature and our previous results we hypothesize that the immune activity against certain tumor-associated antigens is a natural defense mechanism targeting malignant cells, and that the identification of corresponding naturally occurring antibodies could provide a predictive screening tool. Printed glycan array will be used in a separate cohort as previously described (1, 2). For validation purposes, AGA will be measured in an independent cohort established by the same research team, consisting of an equivalent cohort size but from another continent. An Australian cohort was prospectively established and includes over 360 patients with both gynaecological cancers and healthy control patients. The advantage of this validation approach is the fact that both the biobank generation and the PGA experiments are performed by the same personnel in both cohorts.

5.2 Investigation of Specific Glycan Interactions

Aberrant glycosylation of proteins and lipids during malignant transformation results in the appearance of specific glycan structures. These glycans are known as TACA and can be found on cell surfaces and in serum components (3). TACA are therefore promising targets for new diagnostic biomarkers as well as anti-cancer vaccines (4). Altered glycosylation of tumor cells is mainly the consequence of changes in cellular glycosyltransferase activities (5) and is found in almost 100% of cancers, as compared to the expression of oncogenes which is found in less than 30% of cancers (3). Glycan arrays allow determination of the specificities of glycan-binding proteins (GBP) (6), examination of microbiologically relevant glycans (7), study of carbohydrate-processing enzymes (8), and anti-glycan immune responses.

To date, only a few studies investigated a small cohort of glycans and glycoconjugates within a specific biological system (9), by screening AGA within human serum for diagnostic purposes (10, 11) or detecting the immune response to bacterial pathogens (12). In contrast to this, printed glycan array was used to identify specific ovarian cancer AGA patterns. P_1 (*Galα1-4Galβ1-4GlcNAc)* a member of the P blood group system differentiated most significantly between the cancer and the control cohort. The physiological function of the P blood group system in general is still unknown. The detection of AGA to P_1 as a potential

ovarian cancer associated carbohydrate antigen could imply that, similarly to mesothelioma cells (13), ovarian cancer cells contain a P_1 epitope recognized by the human immune system.

5.2.1 Hypothesized Strategy

Glycans presented on eukaryotic cells are known targets for AGA as well as other GBP, such as galectins or siglecs. They are able to specifically recognize and interact with carbohydrates. GBP are involved in biological processes such as immune recognition and regulation, inflammatory responses, cytokine signaling, and cell adhesion. It is therefore important to further investigate and understand the glycobiological and immunological impact behind the identification of P_1 as an ovarian cancer–specific carbohydrate structure with antigenicity.

Our main aim for future studies will therefore be the examination of GBP and glycan-presenting proteins (=glycosylated proteins) in order to understand the biological function of a specific carbohydrate structure in ovarian carcinogenesis. To investigate the role of specific glycans and their immune recognition biomolecules it is necessary to synthesize appropriate carbohydrates of high purity. Synthesized and biologically validated carbohydrates will then be used to affinity purify specific binding partners from pooled ascites, blood serum samples or ovarian cancer cell lines (Figure 25/1). After established affinity purification, eluted GBP will be validated based on their protein content by SDS-PAGE and silver staining. In parallel, Western blot analysis will be used to control the presence of human AGA of immunoglobulin subtype IgA, IgG and IgM. Eluates after affinity purification should contain immunoglobulins and are therefore additional controls for affinity purification. Identification of potential GBP will be performed by liquid chromatography mass spectrometry. The specificity and binding affinity of identified GBP will be studied by printed glycan array as well as custom-made suspension array and ELISA using commercially available mAb. Immuno-blotting will be performed to evaluate identified GBP and to test their presence in applied positive and negative controls.

Glycosylation is an enzymatic process that links various saccharides to produce glycan structures. This product is attached to proteins, lipids and other organic molecules. Glycosylated proteins serve in membranes and in secretions for a variety of structural and functional roles (14). Specific glycans (*e.g.* P_1) have been identified as recognisable by the adaptive innate immune system and can therefore potentially serve as ovarian cancer-associated carbohydrate antigen. If this antibody-glycan interaction is cancer-specific and which underlying glycan-presenting proteins are involved, is yet unclear. In our case, glycan-

presenting proteins are defined as proteins which are post-translationally modified by a specific glycan structures such as P_1.

In a second approach we aim to study glycan-presenting proteins by using nanoparticle-antibody conjugations (15). Two classes of nanoparticles will be used: (A) fluorescent dye-loaded silica and (B) magnetic (Fe_3O_4) nanoparticles in order to localize and purify glycosylated proteins containing glycans of interest. AGA covalently coupled to nanoparticles will be tested on corresponding chemically synthesized Sepharose-PAA-glycan and on cell lines. Purified, validated and fluorescence-nanoparticle labeled AGA will then be used to localize specific glycan structures in paraffin-embedded patient tissues (A). Magnetic-nanoparticle labeled AGA will be used to bind glycosylated proteins (B). Captured proteins will then be purified *via* a magnetic column and eluted prior to structural analysis with mass spectrometry (Figure 25/2-4).

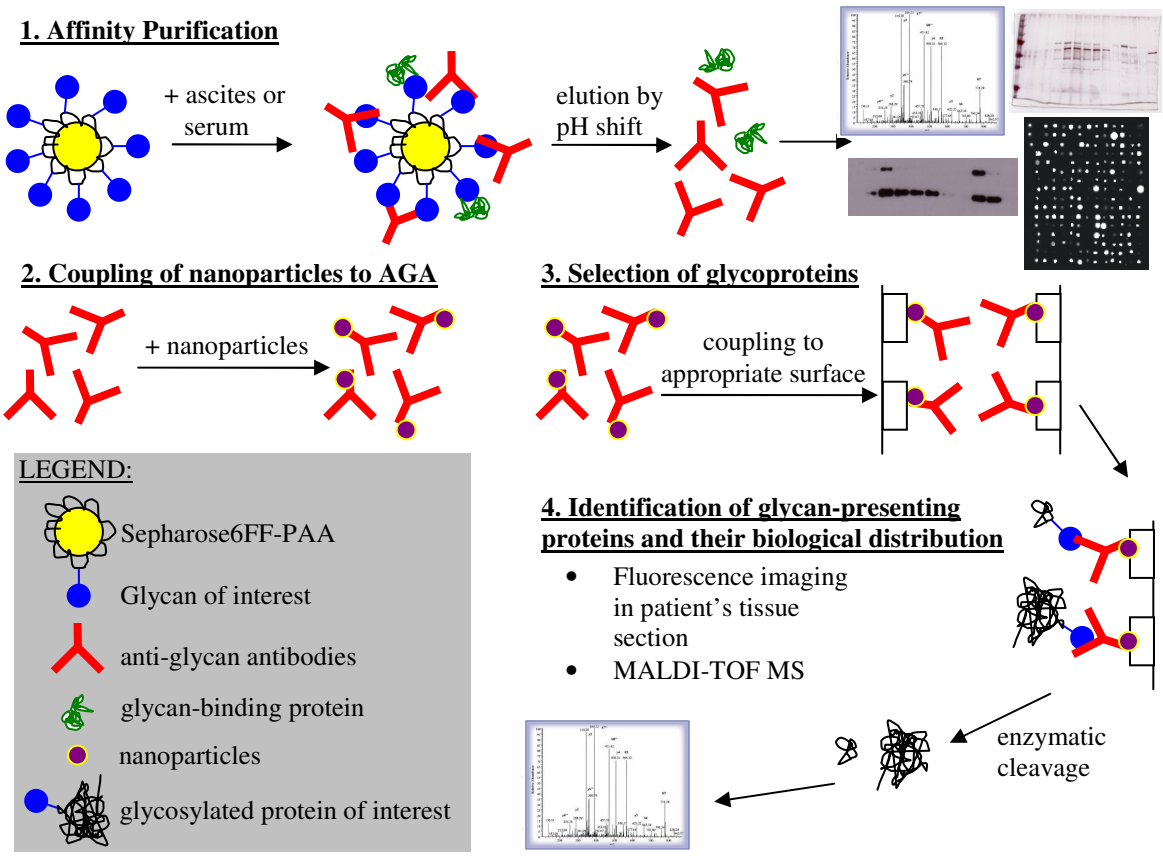


Figure 25 Strategy for identification of ovarian cancer candidate-associated GBP.

5.2.2 Preliminary Data

5.2.2.1 Synthesis of Glycosorbents

Cognate adsorbent (glycosorbents) structures were produced for affinity chromatography Sepharose6FF and coupled to polymer presenting glycans. The synthesized glycosorbents were generated in two major steps: (1) synthesis of polymeric neoglycoconjugates by activated polymer reaction with ω-aminoalkylglycosides (16), and (2) synthesis of final glycosorbents. Chemically synthesized glycans of high purity, end-labeled by a monomeric spacer and terminated by an activated amine group were coupled to polymer polyacrylamide (PAA). After blocking with ethanolamine the reaction was verified by thin layer chromatography followed by Ninhydrin treatment for the detection of free amine groups. Sepharose6FF was added to PAA-glycan at pH 8.5 adjusted by triethanolamine. Incubation with Sepharose6FF-PAA-glycan was blocked by ethanolamine and the reaction verified by thin layer chromatography and phosphorous acid reaction for the detection of carbohydrates.

5.2.2.2 Verification of Synthesized Glycosorbents

As a proof of principle, chemically synthesized carbohydrates were studied in their binding ability towards human antibodies by incubation of glycosorbents containing blood group A trisaccharide (A_{tri}), a minimal AB0 blood group determinant. Pooled serum of 10 healthy controls with blood group 0 was applied to Sepharose6FF-PAA-A_{tri}, and bound human anti-A_{tri} antibodies were visualized using Tetramethyl Rhodamine Iso-Thiocyanate (TRITC) labeled goat anti-human IgA, IgG & IgM (Figure 26). The negative control was not incubated with human serum and showed no signal (Figure 26). Pooled samples of blood group A with expected low or no antibodies to A_{tri} were incubated with Sepharose6FF-PAA-A_{tri}. Low signals were detected with fluorescence microscopy (data not shown), which could also be due to cross-reactivity or unspecific binding of AGA. Verification of synthesized glycosorbents revealed that chemically synthesized carbohydrates are suitable for further functional studies.

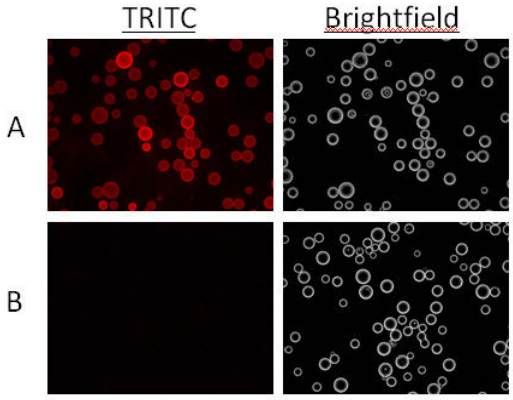


Figure 26 Serum containing anti-A_{tri} antibodies bound to Sepharose6FF-PAA-A_{tri}.
Blood group 0 containing A_{tri} binding antibodies (A) were visualized by goat anti-human IgA, IgG, IgM conjugated TRITC (A/TRITC), in contrast to control without serum (B/TRITC) with no fluorescence signal. (TRITC) Tetramethyl Rhodamine Iso-Thiocyanate labeled goat anti-human IgA, IgG and IgM; Brightfield images at same glass slide position.

5.2.2.3 Verification of P_1- Affinity Purified Eluates

Affinity chromatography was performed to isolate specific GBP. Sepharose conjugated P_1 (*Galα1-4Galβ1-4GlcNAc*-PAA-Sepharose) was selected as first glycan of interest. A pool of ascites from 17 ovarian cancer patients was preprocessed as source of potential P_1- binding proteins. Performing 1D SDS-PAGE in combination with silver staining, we visualized several affinity purified protein bands. These bands appeared in different eluates and ranged from 25kDa to 100kDa in molecular size (Figure 27/1). Additionally, immuno-blotting using anti-human IgA, IgG and IgM biotin streptavidin detection system revealed immunoglobulins bound to P_1 (Figure 27/2). These first experiments using after affinity purification confirmed can be seen as proof of principle and show the presence of AGA as well as other GBP.

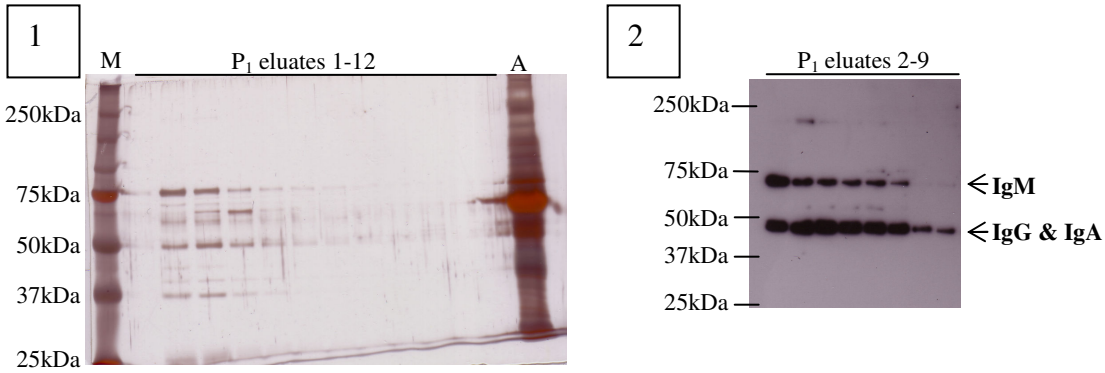


Figure 27 Validation of P_1 affinity purified glycan binding proteins.
Collected eluates of affinity purified samples were separated by SDS-PAGE and proteins visualized by silver nitrate staining procedure (1) along with the standard marker (M; Precision Plus Protein Marker [Bio-Rad]) and positive control 1:500 diluted ascites pool (A). Eluates 2 to 9 separated by SDS-PAGE and AGA of immunoglobulin subtypes are detected by goat anti-human IgG, IgA and IgM biotin-streptavidin system (2).

5.2.2.4 Validation of Affinity Purified Anti- Glycan Antibodies

Following affinity purification we measured the total protein content and presence of human AGA using Western blotting and AGA bound to P_1 using direct ELISA. Cross-reactivity of purified AGA was studied in a self-made ELISA using non-related P_1 glycan structures.

ELISA revealed a high optical density for glycan P_1 and low or no cross-reactivity to other glycans such as AB0 blood group antigens or rhamnose, a glycan whose corresponding antibodies are present in high amounts in serum across all humans (17) (Figure 28, A). As would be expected, low signals were detected in control samples containing unspecific AGA (Sepharose6FF-PAA, without glycan) (Figure 28, B). This indicates a highly specific purification of anti-P_1 antibodies, which is essential for our further experiments. This experiment therefore shows promising results in that AGA bound to Sepharose6FF-PAA-P_1 was highly affinity purified with low cross-reaction to other carbohydrate structures.

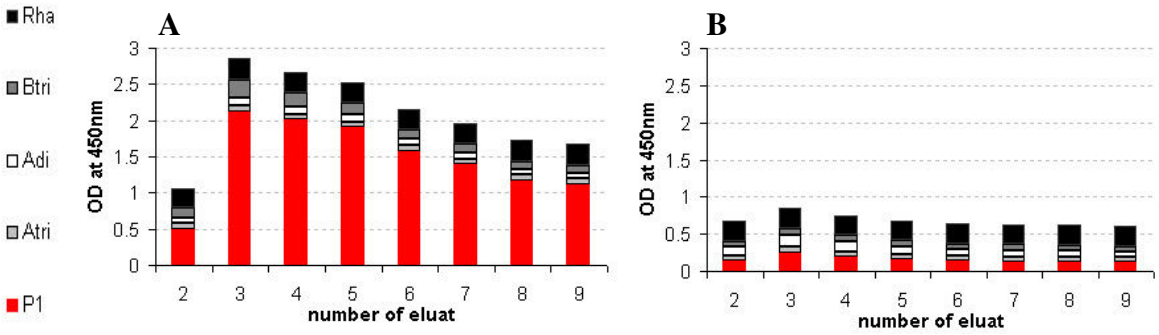


Figure 28 Affinity purified anti-glycan antibodies against P_1 validated by ELISA.
Different eluates (2-9) arising from Sepharose6FF-PAA-P_1 (A) and Sepharose6FF-PAA (B) were subjected to Rha, B_{tri}, A_{di}, A_{tri} and P_1 and are presented as cumulative bar graphs. High specificity of P_1-purified AGA was achieved in eluates 3 to 7 (A) with low signals to other applied glycan structures. Overall low signals were observed in control samples (B). In A, decreasing signal intensity to P_1 refers to increasing volume of elution buffer. Low-AGA binding was observed in eluate 1 and 2.

5.2.2.5 Mass Spectrometry to Identifies Glycan-Binding Proteins

Mass spectrometry was used to identify potentially interacting binding partners to identified glycans; P_1 affinity purified proteins were solubilized for better susceptibility of enzymatic cleavage by RapiGest™ SF Surfactant. Remaining inter- and intra-molecular disulfide bonds were reduced by Tris (2-carboxylethyl) phosphine hydrochloride (TCEP HCl) and samples were alkylated by iodocetamide and digested overnight with trypsin. To avoid contamination, digested peptides were purified with ultramicro spin column and redissolved in 0.1% formic acid for injection into the mass spectrometer.

Liquid chromatography mass spectrometry analysis of P_1-affinity purified samples was performed on a Thermo Fourier Transformed-LTQ mass spectrometer, connected to an electrospray ionizer. Acquired MS scans were searched against the human International Protein Index (IPI) protein database. Final identification revealed 53 potential P_1-binding proteins. Beside a broad spectrum of detected immunoglobulin subtypes, other non-immunoglobulin related proteins were identified (Table 12). Most of the identified proteins are associated to immune response processes. Interestingly, beside expected P_1 related immunoglobulins, known GBP such as Ficolin 2 (a serum lectin p35) and complement-associated protein SP-40 (Vitronectin) were identified. Whilst these both proteins are only known to bind carbohydrate structures, we identified a third protein T-kinin (Kininogen) which is known to be involved in ovarian cancer.

Due to unspecific binding and cross-reactivity, our preliminary experiments revealed that the polymer polyethylenglycol could cause high background signals. The presence of this polymer could be explained by the Tween 20 washing steps which might have washed off loosely or unspecifically bound reactants, or unstable polymer coming from the glycosorbent. The presence of polyethylenglycol could possibly enhance unspecific binding and will be avoided in further experiments. As an additional control, PAA on Sepharose6FF will be used in affinity chromatography to identify unspecific binding proteins. Nevertheless, our preliminary data of P_1 affinity purified ascites samples revealed specific AGA and GBP which are known to be associated to ovarian cancer as well as glycan-binding activity.

Uniprot accession	**Gene Symbol**	**Protein Description**	**protein probability**	**sequence coverage [%]**	**number unique peptides**	**total number peptides**
P01024	C3	Complement C3 precursor	1	37.6	72	92
P01031	C5	Complement C5 precursor	1	8	13	13
P13671	C6	Complement component C6 precursor	1	9.9	9	9
P10643	C7	Complement component C7 precursor	1	5.5	4	4
P00739	HPR	Haptoglobin-related protein precursor	1	27.9	6	6
P22692	IGFBP4	Insulin-like growth factor binding protein 4 precursor	1	23.3	4	4
P01042	**KNG1**	**Kininogen** precursor	1	19.9	9	9
P04004	**VTN**	**Vitronectin** precursor	1	6.7	3	3

P01028	C4	Complement C4 precursor	1	1.1	2	2
O43866	CD5L	CD5 antigen-like precursor	1	27.7	9	11
P27918	CFP	Properdin precursor	1	27.1	10	12
P00751	CFB	Complement factor B precursor	1	23	19	19
Q15485	**FCN2**	**Ficolin 2** precursor	0.9985	6.4	2	2
P02765	AHSG	Alpha-2-HS-glycoprotein precursor	0.9913	5.4	1	1
P61769	B2M	Beta-2-microglobulin precursor	0.9913	8.4	1	1
Q02413	DSG1	Desmoglein 1 precursor	0.987	1.3	1	1
P10909	CLU	Clusterin precursor	0.9819	3.6	1	1
P00747	PLG	Plasminogen precursor	0.9817	3	2	2
P49908	SEPP1	Selenoprotein P precursor	0.976	2.6	1	1
Q9BXR6	CFHR5	Complement factor H-related protein 5 precursor	0.976	1.8	1	1
P01034	CST3	Cystatin C precursor	0.9752	6.2	1	1
P04003	C4BPA	C4b-binding protein alpha chain precursor	0.9743	1.7	1	1
P81605	DCD	Dermcidin precursor	0.9522	10	1	1
P02753	RBP4	Plasma retinol-binding protein precursor	0.9333	5.5	1	1
P61626	LYZ	Lysozyme C, type P precursor	0.9401	9.2	1	1

Table 12 P_1 binding proteins identified *via* mass spectrometry.
P_1-binding proteins enriched by affinity chromatography revealed 53 P_1-binding partners. The first 25 potential GBP are listed, in bold are identified proteins with known glycan binding activity or ovarian cancer association.

5.3 References

1. Blixt O*, et al.* (2004) Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. *Proc Natl Acad Sci U S A* 101(49):17033-17038.
2. Huflejt ME*, et al.* (2009) Anti-carbohydrate antibodies of normal sera: findings, surprises and challenges. *Mol Immunol* 46(15):3037-3049.
3. Hakomori S (2001) Tumor-associated carbohydrate antigens defining tumor malignancy: basis for development of anti-cancer vaccines. *Adv Exp Med Biol* 491:369-402.
4. Fuster MM & Esko JD (2005) The sweet and sour of cancer: glycans as novel therapeutic targets. *Nat Rev Cancer* 5(7):526-542.
5. Chandrasekaran EV, Xue J, Neelamegham S, & Matta KL (2006) The pattern of glycosyl- and sulfotransferase activities in cancer cell lines: a predictor of individual cancer-associated distinct carbohydrate structures for the structural identification of signature glycans. *Carbohydr Res* 341(8):983-994.

6. Moller I, *et al.* (2008) High-throughput screening of monoclonal antibodies against plant cell wall glycans by hierarchical clustering of their carbohydrate microarray binding profiles. *Glycoconj J* 25(1):37-48.
7. Disney MD & Seeberger PH (2004) The use of carbohydrate microarrays to study carbohydrate-cell interactions and to detect pathogens. *Chem Biol* 11(12):1701-1707.
8. Park S & Shin I (2007) Carbohydrate microarrays for assaying galactosyltransferase activity. *Org Lett* 9(9):1675-1678.
9. Krishnamoorthy L & Mahal LK (2009) Glycomic analysis: an array of technologies. *ACS Chem Biol* 4(9):715-732.
10. Dotan N, Altstock RT, Schwarz M, & Dukler A (2006) Anti-glycan antibodies as biomarkers for diagnosis and prognosis. *Lupus* 15(7):442-450.
11. Wandall HH, *et al.* (2010) Cancer biomarkers defined by autoantibody signatures to aberrant O-glycopeptide epitopes. *Cancer Res* 70(4):1306-1313.
12. Bovin NV & Huflejt ME (2008) Unlimited glycochip. *Trends Glycosci. Glycotechnol.* 20(115):245-258.
13. Spitalnik SL, *et al.* (1986) Glycolipid antigen expression in human lung cancer. *Cancer Res* 46(9):4751-4755.
14. Varki A, *et al.* (1999) *Essentials of Glycobiology* (Cold Spring Harbor Labartory Press, Plainview (NY)) p p.41.
15. Thierry B, Al-Ejeh F, Brown M, Majewski P, & Griesser HJ (2009) Immunotargeting of advanced functional nanostructures for cancer diagnosis and treatment. *Advanced Materials* 21:541-545.
16. Bovin NV, *et al.* (1993) Synthesis of polymeric neoglycoconjugates based on N-substituted polyacrylamides. *Glycoconj J* 10(2):142-151.
17. Dotan I, *et al.* (2006) Antibodies against laminaribioside and chitobioside are novel serologic markers in Crohn's disease. *Gastroenterology* 131(2):366-378.

6 Summary

Epithelial ovarian cancer (EOC) has the highest mortality in gynaecological cancers with a 5-year survival of only 20% due to detection of 2/3 of cases in advanced stage disease. Based on proteogenomic studies during the last 10 years a paradigm shift has taken place with the increasing insight that this is a biologically heterogeneous disease. The identification of type-specific genetic patterns or protein markers will contribute to the identification of disease origin, and allow early detection, prediction and targeted treatment. We performed the first comprehensive overview on all published studies using functional proteogenomics in EOC, to critically evaluate findings including genetically heterogenous subtypes of EOC, and to develop research questions not yet been addressed. Between 1999 and 2007, 237 studies were published as revealed by systematic online database screening, including Pub Med, Cochrane library and Medline. Most studies were using transcriptomic approaches (63.7%), followed by proteomics (20.3%) and other methods. Although a variety of research questions were addressed, most studies concentrated on the identification of biomarkers for early detection, independent of platform used. Later studies seemed to realize the technical limitations and therefore concentrated rather on molecular patterns to distinguish treatments, drug resistance, metastases and technical microarray limitations. As a result of our summary we found that the main limiting factors remain the low reproducibility and the heterogeneity, mainly of samples, array platforms, bioinformatical analyses, and the disease of EOC in total.

The decision was therefore made to study a more homogeneous group of non mucinous EOC and to use a robust microarray technology which concentrates rather on post-translational modifications, an approach which has not been performed so far. Whilst studies on high-throughput profiling in the field of glycobiology have been published in small numbers, no study on EOC profiling for specific anti-glycan antibodies (AGA) has been performed previously. AGA incorporate a vast and yet insufficiently investigated subpopulation of naturally occurring and adaptive antibodies in humans. Glycan-antibody interactions mediate many important biological processes, such as innate immune responses, recognition of malignant transformation, autoimmune diseases, neurological disorders and host-*versus*-graft rejections. Due to the complexity and heterogeneity of glycan structures, as well as peculiarities of antibody-glycan binding *in vivo*, many important aspects of these interactions are still poorly understood.

Our main aim was to investigate the potential of AGA in the diagnosis of well-defined non-mucinous EOC using a novel printed glycan array (PGA) discovery approach. PGA is the

first high-throughput glycan array containing as many as 203 glycan structures, synthesized to a purity of approximately 98%. We found high reproducibility in measuring AGA which also showed a signal towards blood groups as known from literature. Significant discrimination between healthy controls and cancer samples was found in 24 glycans, with P_1 (*Galα1-4Galβ1-4GlcNAcβ*; $p<0.001$) being the most significant candidate. Higher sensitivity and specificity than currently applied biomarker CA125 was achieved by a panel of multivariate selected and linear combined AGA signals (83.3% and 84.8%, respectively).

To understand the biological properties of AGA it is necessary to study their binding properties. We therefore performed cluster analysis and found AGA which specifically bound to core glycan structures reflected by 21 unique clusters, including *N*- and *O*-linked glycans. To study AGA bindings in greater depth, we developed two compatible 'glycan-based immunoassays' (multiplex suspension array, ELISA), where the same glycans were presented chemically and spherically in different forms. Whilst screening for blood group AB0 antigens A and B trisaccharide, all three assays revealed the same reproducibility and results, in contrast to PGA. Based on these findings we then investigated a previously PGA identified ovarian cancer glycan P_1. Interestingly, among all three methods the overall AGA binding pattern decreased in this setting. PGA showed a significant discriminative power between cancer and healthy serum samples in the patient cohort which the other two methods did not seem to be able to reproduce. This could be influenced by differences in the various assay conditions like different concentrations of glycans as well as serum dilutions.

Using a PGA discovery approach, we detected AGA which could detect non mucinous EOC with a higher sensitivity and specificity than the current tumor marker CA125. These findings indicate that 'glycoarrays' have a large potential for the development of a new generation of biomarkers in EOC. It also demonstrates that tumor-associated glycans are recognized by the immune system. To investigate the glycobiological role of EOC-associated glycans, we are aiming to next purify detected AGA and to validate and quantify them further. Mass spectrometry will be used to identify AGA subtypes as well as potential glycan-binding proteins, which will promote our understanding of role of the identified antibodies.

7 Zusammenfassung

Das epitheliale Ovarialkarzinom (EOC) ist die fünfthäufigste maligne Erkrankung bei der Frau weltweit. Innerhalb der gynäkologischen Krebserkrankungen zeigt es die höchste Sterblichkeit. Dies beruht darauf, dass 75% der Patientinnen erst in einem fortgeschrittenen Stadium diagnostiziert werden und somit die 5-Jahres-Überlebensrate auf nur 20% sinkt. Innerhalb der letzten 10 Jahre kam es aufgrund proteogenomischer Studien zu einem Paradigmenwechsel, so dass EOC heute als biologisch heterogenes Krankheitsbild verstanden wird. Zur besseren Früherkennung, Prognose und gezielten Behandlung ist es von großer Bedeutung, spezifische molekulare Marker zu identifizieren.

In der vorliegenden Arbeit wurden alle bisherigen Publikationen, die eine Analyse des Genoms oder Proteoms beim EOC zum Inhalt hatten, analysiert. Ziel war es herauszufinden, welche Fragestellungen mit welchen Forschungsplattformen und klinischen Proben durchgeführt wurden. In dem Zeitraum von 1999 bis 2007 konnten anhand einer systematischen Online-Datenbank-Analyse in PubMed, Medline und Cochrane 237 Studien identifiziert werden. Die meisten Studien untersuchten das menschliche Transkriptom (63,7%), nur 20.3% hingegen das Proteom. Trotz einer Vielzahl von Fragestellungen beschränkten sich frühe Studien auf die Identifizierung von Biomarkern zur Früherkennung, unabhängig von der verwendeten Plattform. Spätere Studien hingegen konzentrierten sich auf die Untersuchung technischer Grenzen, molekularer Muster, Behandlungsunterschiede, Tumorresistenzen und Metastasen. Zusammenfassend konnte somit gezeigt werden, dass die geringe Reproduzierbarkeit und Heterogenität der Proben, sowie verschiedene Array-Plattformen, bioinformatische Analysen, und das heterogene Krankheitsbild die größte Limitierung für die Tumormarker-Entwicklung darstellen. Eine geringe Anzahl von Studien verwendete Hochdurchsatzmethoden im Bereich der Glykobiologie, jedoch gab es keine solche Studie beim EOC. Zudem wurde bislang keine Studie publiziert, welche spezifische Anti-Glykan Antikörper (AGA) untersucht, obgleich diese eine große und noch unzureichend untersuchte Subpopulation von natürlich vorkommenden und adaptiven Antikörper beim Menschen darstellt. Aufgrund der Analyse bisher publizierter Studien beim EOC wurde in der vorliegenden Arbeit ein Ansatz verfolgt, welcher eine homogene Gruppe von nicht muzinösen EOC und gesunden Kontrollen mit einer robusten Microarray- Technologie kombinierte.

Glykan-Antikörper-Interaktionen vermitteln eine Vielzahl biologischer Prozesse, welche während der Immunantwort, Erkennung der malignen Transformation, Autoimmunerkrankungen, neurologischen Störungen und Graft *versus* Host Disease

(Transplantat-Abstoßung) von Bedeutung sind. Aufgrund der Komplexität und Heterogenität der Glykanstrukturen, sowie Besonderheiten der Antikörper-Bindung *in vivo*, sind diese Interaktionen noch weitestgehend ungeklärt. Ziel der vorliegenden Arbeit war es, das Potenzial der AGA im Blutserum bei der Diagnose einer homogenen Gruppe von EOC mit Hilfe eines neuartigen Printed Glycan Array (PGA)-Ansatzes zu untersuchen. Der PGA umfasst 203 komplexe Glykanstrukturen welche mit einer Reinheit von 98% chemisch synthetisiert wurden. Die Resultate zeigten eine hohe Reproduzierbarkeit der AGA-Signale, welche auch mit dem bekannten AB0- Blutgruppen-System übereinstimmten. Signifikante Unterschiede zwischen gesunden Kontrollen und Karzinom-Patientinnen konnten bei 24 Glykan-assoziierten AGA Signalen mit dem Topkandidaten P_1 (*Galα1-4Galβ1-4GlcNAcβ*; $p < 0{,}001$) identifiziert werden. Eine höhere Sensitivität und Spezifität im Vergleich zum heute verwendeten EOC Biomarker CA125 wurde durch eine multivariate Kombination der AGA-Signale (83.3% bzw. 84.8%) erreicht.

Um ein besseres Verständnis der biologischen Eigenschaften der AGA zu erhalten, wurden Antikörper-Glykan-Bindungseigenschaften genauer untersucht. Eine Clusteranalyse mit den PGA Daten wurde durchgeführt, wobei sich herausstellte, dass AGA spezielle Kern-Glykanstrukturen binden, welche sich in 21 einzigartigen Clustern widerspiegeln, einschließlich *N*- und *O*- Glykosylierungen.

Zur weiteren Untersuchung der AGA-Bindungseigenschaften wurden zwei kompatible "Glykan-basierte Immunoassays" entwickelt. Alle drei Methoden (Multiplex Suspension Array, ELISA, PGA) verwenden hierbei die gleichen chemisch synthetisierten Glykane, präsentieren diese jedoch sphärisch in verschiedenen Formen. Die Untersuchung von AB0 assoziierten A- und B-Trisacchariden zeigte in allen drei Methoden ähnliche Ergebnisse. Basierend auf diesen Resultaten wurde eine weitere Analyse des zuvor identifizierten Top-Kandidaten P_1 durchgeführt. Hierbei stellte sich heraus, dass die AGA Bindungsmuster in ihrer Konkordanz reduziert waren. Der PGA zeigte das deutlichste Diskriminierungs-Potential zwischen dem Serum Gesunder und Karzinom-Patientinnen. Dies könnte durch unterschiedliche Testbedingungen, wie verschiedene Glykankonzentrationen oder Serumverdünnungen beeinflusst worden sein.

Spezifische AGA für nicht muzinöse EOC erreichten eine höhere Sensitivität und Spezifität als der aktuelle Tumormarker CA125. Neben dem Potenzial für eine neue Generation von Tumormarkern konnte weiterhin gezeigt werden, dass tumor-assoziierte Glykane vom Immunsystem erkannt werden.

In weiterführenden Untersuchungen soll die Rolle glykobiologischer, EOC-assoziierter Glykane durch Aufreinigung von AGA, deren Validierung und Quantifizierung zugrunde gelegt werden. Des Weiteren sollen Massen-spektrometrische Methoden zur Identifizierung von AGA Subtypen sowie potenziellen Glykan-bindenden Proteinen eingesetzt werden, um das Verständnis ihrer biologischen Bedeutung zu verbessern.

8 Contributions

8.1 Manuscript published in Biomarkers in Med (2009) 3(6), 743-756

The special report published in Biomarker in Medicine 2009 with the title "**Proteogenomic studies in epithelial ovarian cancer: established knowledge and future needs**" summarizes important findings based on high-throughput technologies in the field of epithelial ovarian cancer from 1999 to 2007. This project was designed and given to me by my supervisor Dr. Viola Heinzelmann-Schwarz. The systematic literature search was performed completely by Francis Jacob. This includes the development of a strategy and search machine for publications as well as data analysis. The development of aims and the analysis of the data generated was performed by Francis Jacob under supervision of Dr. Viola Heinzelmann-Schwarz. The OLOV database was previously designed by Dr. Matthias Heinzelmann and published in a publication (*Heinzelmann-Schwarz V et al, Cancer Res 2004*). In this study, however, we updated OLOV with up-regulated genes within all identified transcriptomic approaches, hereby highlighting the most promising ovarian cancer-associated genes/ proteins. Further support for this study was given from the statistical point of view by our collaborator Dr. Darlene Goldstein.

8.2 Manuscript in preparation (1)

The second manuscript entitled "**Serum anti-glycan antibody detection of non-mucinous ovarian cancers using printed glycan array**" represents the main part of this PhD thesis. This study was designed by the supervisor Dr. Viola Heinzelmann-Schwarz, including the setup of a large biobank, national and international collaborations and financial grant acquisition for this project. The main aim was the use of a new technology (printed glycan array), available *via* a collaboration with the Scripps Institute in San Diego to study the technical and biological distribution of anti-glycan antibodies in serum of healthy and ovarian cancer patients. The whole experimental procedure included the establishment of a large biobank after ethical approval (performed in parts by Francis Jacob), the experiments around the printed glycan array and the bioinformatical analysis (both parts performed by Francis Jacob under supervision). The printed glycan array experiments included (1) in-house printing of custom glycan arrays, (2) experimental development and screening of human serum samples and (3) mathematical data pre processing. All three steps were performed by Francis Jacob under guidance of the San Diego team Dr. Margaret Huflejt and Prof. Marko Vuscovic. The use of technical and biological relevant glycans was suggested by Francis Jacob.

Bioinformatical data analysis merging the clinical bio-databank PEROV (designed by Dr. M. Heinzelmann) to the printed glycan array data was performed by Francis Jacob under direct supervision of Dr. Darlene Goldstein. The whole project from its initial planning phase to the written manuscript was managed and supervised by Dr. Viola-Heinzelmann Schwarz.

8.3 **Manuscript in preparation (2)**

The third manuscript presented in this PhD thesis entitled "**Comparison of three glycan-based immunoassays: suspension array, ELISA and printed glycan array in the detection of human anti-glycan antibodies**" describes three in-house developed glycan-based assays. In this publication we aim to investigate methodological as well as biochemical dependencies of glycan presentation to human anti-glycan antibodies. Results previously generated by printed glycan array (manuscript 2) were compared to two independent glycan-based assays which were developed within the Translational Research Group (TRG) at the University Hospital Zurich; Dr. Viola Heinzelmann-Schwarz). Bead-based suspension array as a technology was primarily developed by Dr. Tatiana Pochechueva whilst still being a Post-doctoral fellow in Prof. Nicolai Bovin's laboratory in Moscow. The specific suspension array which is described within this paper, however, was developed by Dr. Pochechueva whilst being a Post-doctoral Fellow under Dr. Viola Heinzelmann-Schwarz. She was supported in these experiments by Francis Jacob. The third method, glycan-based ELISA was established and the experiments performed by Francis Jacob. Human serum samples of benign and malignant gynaecological cases were selected and anti-glycan antibody levels measured and analysed per each glycan-based assay. As biological validation AB0-blood group related A and B trisaccharides were selected and anti-glycan antibody binding studied. In addition, the top candidate from publication 2, trisaccharide P_1, was selected and profiled in the same cohort. Clinicopathological extracts from the PEROV database (Dr. Matthias Heinzelmann) were correlated to the experimental findings and statistical analysis performed by Francis Jacob under supervision of Dr. Darlene Goldstein. The whole project was designed, planned and supervised from the financial acquisition till the manuscript preparation by Dr. Viola Heinzelmann-Schwarz. Prof. Nicolai Bovin has supervised partially the suspension array part and has been involved in the manuscript writing due to his expertise in this field.

9 Acknowledgement

I would like to thank all the people who participated in the success of this work, in particular:

- I am most grateful to my wife Anett for her continuous and unquestioned support during the last years. Apart from supporting me in my work I am grateful for Anett giving birth to our beautiful baby Paula who was born on 6th November 2009. Parenting has certainly opened my eyes to absolute happiness. My parents, Petra and Wolfgang Jacob and my parents in-law Monika and Klaus-Dieter Brümmer who always believed in me and thus made this work possible.
- PD Dr. Viola Heinzelmann-Schwarz, my supervisor, who supported my work patiently; her experience and knowledge were invaluable. Her support, trust, and continued belief in me were essential in my personal and scientific growth. All the members of the Translational Research Group, University Hospital Zürich for their support and the pleasant working atmosphere, particularly Dr. Tatiana Pochechueva for the inspiring discussions about glycobiology and Dr. Matthias Heinzelmann for his endless support of my IT needs. All the members of the Gynecology Research Department, particularly PD. Dr. André Fedier, Head of the Research Department of Gynaecology.
- Members of Cellexicon, Inc., La Jolla, CA, USA for sharing their knowledge of printed glycan array technology, particularly Dr. Margaret E. Huflejt for the in-sight into glycobiology and immunology and Prof. Marko Vuskovic who tutored me in pattern recognition and data pre-processing.
- Dr. Darlene R. Goldstein (EPFL Lausanne) who taught me bioinformatical analyses including the statistical analysis using R.
- All the members of the Laboratory of Carbohydrate Chemistry, Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow, Russia for their support and fruitful input. In particular, I would like to thank Prof. Nicolai Bovin who supported me with ideas, critical comments and technical equipment at all times. Without the work of his laboratory this work would not be possible.
- For external supervision by the members of my PhD committee Prof. Christian von Mering (University of Zurich) and Prof. Josef Jiricny (University of Zurich).
- All collaborators: Department of Systems Biology, ETHZ, Prof. Rudi Aebersold, Ruth Hüttenhain; Institute of Clinical Pathology, University Hospital Zurich, Prof. Holger Moch, Prof. Rosmarie Caduf; Department of Gynaecology, University Hospital Zurich, Prof. Daniel Fink.

- All members and PhD students at the Cancer Biology PhD program (ETH Zurich/ University of Zurich)

This work was supported by the Cancer League of the Canton of Zurich; Oncosuisse; Swiss National Foundation, European Science Foundation; Cellexicon, Inc. La Jolla, CA, USA; National Cancer Institute Grant, USA.

10 Appendix

10.1 Abbreviations

®	registered trade mark
µl	microliter
µM	micromolar
2D PAGE	Two-Dimensional Poly-Acrylamid Gel Electrophoresis
AGA	Anti-Glycan Antibodies
ANN	Neural Network Analysis
Asn	asparagine
A_{tri}	*GalNAcα1-3(Fucα1-2)Galβ*
AUC	Area under curve
BCG	bacille Calmette-Guérin
BG	blood group
BL	borderline
BMP7	bone morphogenetic protein 7
BRAF	v-raf murine sarcoma viral oncogene homolog B1
BRCA1	gene breast cancer 1, early onset
BRCA2	gene breast cancer 2, early onset
BSA	Bovin Serum Albumin
B_{tri}	*Galα1-3(Fucα1-2)Galβ-*
CA125	mucin 16 (MUC16), cell surface associated
CA15-3	Carbohydrate Antigen 15-3
CA19-9	Carbohydrate Antigen 19-9
CA72-4	Cancer Antigen 72-4
CD43	leukosialin
$CD5^+$	Lymphocyte antigen T1/Leu-1
cDNA	complementary DNA
CFG	Consortium for Functional Glycomics
CFR-1	cystein-rich fibroblast growth factor receptor
CGH	comparative genomic hybridization
ClCOC	clear cell ovarian cancers
cm	centimeter
CTNNB1	catenin (cadherin-associated protein), beta 1, 88kDa
CV	Coefficient of Variation
DDR1	discoidin domain receptor tyrosine kinase 1
DMH	differential methylation hybridization
DNA	deoxyribonucleic acid
EGFR	epidermal growth factor receptor
ELISA	Enzyme-Linked Immunosorbent Assay
EM	extracellular matrix
EMA	epithelial membrane antigen
EnOC	endometrioid ovarian cancers
EOC	epithelial ovarian cancers
ERBB2	v-erb-b2 erythroblastic leukemia viral oncogene homolog 2
ESI	electrospray ionization
FA2	agalactosyl biantennary glycan
FIGO	International Federation of Gynecology and Obstetrics
Fuc	fucose
***G**al*	galactose
GalNAc	*N*-acetyl galactosamine
GALNT1	UDP-*N*-acetyl-alpha-D-galactosamine:polypeptide N acetylgalactosaminyl transferase 1
GAS6	growth arrest-specific 6
GATA1	GATA binding protein 1 (globin transcription factor 1)
GATA2	GATA binding protein 2 (endothelial transcription factor GATA 2)
GBP	Glycan-Binding Proteins
GD3	*Neu5Aca2-8Neu5Aca2-3Galβ1-4Glcβ*
Glc	glucose
GlcA	glucuronic acid
GlcNAc	*N*-acetyl glucosamine
Globo H	*Fucα1-2Galβ1-3GalNAcβ1-3Galα1-4Galβ1-4Glcβ*
Glyc	Glycan
GM2	*Neu5Aca2-3(GalNAcβ1-4)Galβ1-4Glcβ*
GPI	glycophosphatidylinositol
GRP78	glucose-regulated protein, 78kDa
GT3	*Neu5Acα2-8Neu5Acα2-8Neu5Acα2-3Galβ1-4Glc*
H type 1	*Fucα1-2Galβ1-3GlcNAc*
H type 2	*Fucα1-2Galβ1-4GlcNAc*
HE4	WAP four-disulfide core domain 2 (*alias* WFDC2, WAP5)
HGSC	high-grade serous carcinoma
HMFG-IIIC12	milk fat globule-EGF factor 8 protein
HOXA7	homeobox A7
HOXB7	homeobox B7
HRPO	horseradish peroxidase
***I**dA*	iduronic acid
IgA	Immunoglobulin A
IgD	Immunoglobulin D
IgG	Immunoglobulin G
IgM	Immunoglobulin M
IL-8	interleukin 8
IMAC	immobilized metal affinity chromatography
INHA	inhibin, alpha
INHB	inhibin, beta
kDa	kilodaltons
KLK8	kallikrein-related peptidase 8
KRAS	v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog
LASA	lipid-associated sialic acid
LC	liquid chromatography

Le^b	*Fucα1-4(Fucα1-2Galβ1-3)GlcNAcβ*
Le^x	*Galβ1-4(Fucα1-3)GlcNAβ*
Le^y	*Fucα1-3(Fucα1-2Galβ1-4)GlcNAcβ*
LGSC	low-grade serous carcinoma
LMP	low-malignant potential
L-PHA	leukocytic phytohemagglutinin
mAb	monoclonal antibody
MALDI	matrix-assisted laser desorption/ionization
Man	mannose
MCF-7	Michigan Cancer Foundation-7 cell line
medTSI	median total signal intensity
MFI	median fluorescence intensity
MGED	Microarray Gene Expression Data
MIAME	Minimal Information About Microarray Experiments
MIAPE	Minimal Information About Proteomics Experiments
MISFISHIE	Minimum Information Specification For In Situ Hybridization and Immunohistochemistry Experiments
ml	milliliter
MOC	mucinous ovarian cancers
MS	Mass Spectrometry
MSLN	mesothelin
MUC1	mucin 1, cell surface associated
n	number
Neu5Ac	neuraminic acid
Neu5Gc	*N*-glyconeuraminic acid
NHS	*N*-hydroxysuccinimide
NSAID	non-steroidal anti-inflammatory drugs
OD	optical density
OI4	collagen, type I, alpha 2; osteogenesis imperfecta type IV
OLOV	OverLapping OVarian data base
OSE	ovarian surface epithelium
OVCAR-3	human ovarian cancer cell line 3
OVXI	mAb recognizes modified Lewis X determinant on mucin
$\mathbf{P}_1$	*Galα1-4Galβ1-4GlcNAc*
p53	gene for tumor protein p53
PAA	polyacrylamide
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PEROV	Peritoneal and Ovarian Cancer data base
PGA	printed glycan array
PIK3CA	phosphoinositide-3-kinase, catalytic, alpha polypeptide
Pk	*Galα1-4Galβ1-4Glc*
PTEN	phosphatase and tensin homolog
PTM	post-translational modification
QS-21	Quillaja Saponaria 21
RFU	relative fluorescence unit
Rha	Rhamnose
RMI	Risk of Malignancy Index
RNA	ribonucleic acid
ROC	Receiver Operating Characteristics
S100A7	S100 calcium binding protein A7
SA	Suspension Array
SAX2	Strong Anion Exchanger 2
SELDI	Surface-Enhanced Laser Desorption and Ionization
SEM	scanning electron microscopy
Ser	serine
SEREX	serological analysis of recombinant cDNA expression
sLe^a	*Neu5Acα2-3Galβ1-3(Fucα1-4)GlcNAcβ*
sLe^x	*Neu5Acα2-3Galβ1-4(Fucα1-3)GlcNAβ*
SNP	single nucleotide polymorphism
SOC	serous ovarian cancers
SSEA3	*Galβ1-3GalNAcβ1-3Galα1-4Galβ1-4Glcβ1-1*Ceramide
sTn	*Neu5Acα2-6GalNAcα*
T antigen	Thomson-Friedenreich antigen
TACA	tumor-associated carbohydrate antigen
TCEP HCl	Tris (2-carboxylethyl) phosphine hydrochloride
TCOC	transitional cell ovarian cancer
TF	Thomson-Friedenreich antigen
TFCP2	transcription factor CP2
TGF-beta	transforming growth factor-beta
Thr	threonine
TMB	3,3′,5,5′-Tetramethylbenzidine
Tn	*GalNAcα*
TOF	Time of Flight
TRITC	Tetramethyl Rhodamine Iso-Thiocyanate
U	Units
v	version
VEGF	vascular endothelial growth factor
VTCN1	V-set domain containing T cell activation inhibitor (*alias* B7-H4)
WCX2	Weak Cation Exchanger 2
WHO	World Health Organization
X*yl*	xylose

10.2 List of Figures

10.3 List of Tables

10.4 **Curriculum vitae (deutsch)**

JACOB, Francis

Geburtsdatum:	30. Juni 1980
Heimatort:	Pirna, Sachsen, Deutschland
	verheiratet mit Anett Jacob geb. Brümmer
	ein Kind: Paula Jacob, geboren am 06.11.2009

Ausbildung

Gymnasium

Gymnasium	Johann-Gottfried Herder Gymnasium, Pirna, Deutschland
Abschlussjahr	1999
Abschlussnote	2.1 (Zeugnis der allg. Hochschulreife; 1 (sehr gut) bis 6 (ungenügend))

Studium

Universität	Technische Universität Dresden, Dresden, Deutschland
Fakultät	Fakultät Mathematik und Naturwissenschaften
Studiendauer	10. Semester (regulär)
Abschlussdatum	09.11.2005
Diplomarbeit	„ Untersuchungen zur Involvierung der Apyrasen im Pollenschlauchwachstum und in Glykosylierungen einhergehend mit der Entwicklung einer *AtApy2-cMyc*-Fusion“
Abschlussnote	1.7 (Diplombiologe; 1 (sehr gut) bis 6 (ungenügend))

Dissertation

Anstellung als Doktorand	01.09.2006 bis 30.06.2010 (Universität Zürich)
Doktoranden- Pogramm	Cancer Biology PhD Program, Universität und Eidgenössisch-Technische Hochschule Zürich

10.5 **Publications and Conference Contributions**

Peer-Reviewed Articles

Wu, J., Steinebrunner, I., Sun, Y., Butterfield, T., Torres, J., Arnold, D., Gonzalez, A., **Jacob, F.**, Reichler, S., Roux, S.J. (2007) “Apyrases (nucleoside triphosphate-diphosphohydrolases) play a key role in growth control in *Arabidopsis*.” Plant Physiol 144(2): 961-975.

Jacob, F., Goldstein, D.R., Fink, D., Heinzelmann-Schwarz, V. (2009) "Proteogenomic studies in epithelial ovarian cancer: established knowledge and future needs." Biomarkers Med. 3(6): 743-756

Manuscript submitted or in preparation

Jacob F., Goldstein D.R., Bovin N.V., Pochechueva T., Spengler M., Caduff R., Fink D., Vuskovic M.I., Huflejt M.E., Viola Heinzelmann-Schwarz; "Serum anti-glycan antibody detection of non-mucinous ovarian cancers using printed glycan array."

Jacob F., Pochechueva T., Goldstein D.R., Huflejt M.E., Chinarev A., Spengler M., Fink D., Bovin N.V., Heinzelmann-Schwarz V.; "Detection of human anti-glycan antibodies by three different immunoassays: antigen-dependent similarities and disagreements."

Conference presentations

Published Abstracts

Meier, M., **Jacob, F.**, Streich, M. Heinzelmann, M., Fink, D. Heinzelmann-Schwarz, V. (2007) Distinct Clinical Entities within Ovarian Cancer Patients as Detected by Functional Genomics – an Epidemiologic Approach. *Gynaekol. Geburtshilfliche Rundsch* 47, 182.

Jacob, F., Fink, D., Heinzelmann-Schwarz, V. (2007) Seven Years of Functional Genomics in Ovarian Cancer: Current Standard of Knowledge and Future Directions. *Gynaekol. Geburtshilfliche Rundsch* 47, 186.

Ukegjini, K., **Jacob, F.**, Fink, D., Heinzelmann-Schwarz, V. (2007) Secreted Frizzled Receptor Protein 4 Is Detectable in Serum and Ascites of Ovarian Cancer Patients and Is a Marker of Recurrence. *Gynaekol. Geburtshilfliche Rundsch* 47, 188.

Jacob, F., Heinzelmann-Schwarz, V., Vuskovic, M., Chambers, M., Chung, G., Manalili, D., Caduff, R., Fink, D., Huflejt, M. E. (2008) Anti-Glycan Antibodies in Ascites and Serum of Patients with Serous Pelvic Mass. *Proc. Suppl. AACR: Late-Breaking Abstracts*, 39.

Bühler, M., Goldstein, D., **Jacob, F.**, Caduff, R. Fink, D. Heinzelmann-Schwarz, V. (2008) Meta-Analysis reveals growth arrest specific gene 6 as an ovarian cancer specific marker which predicts patient survival. *Gynaekol. Geburtshilfliche Rundsch*

Jacob, F., Caduff, R., Huflejt, M.E., Vuskovic, M. Chambers, J., Chung, G., Manalili, D., Fink, D., Heinzelmann-Schwarz, V. (2008). Patients with serous ovarian cancer produce

equivalent levels of anti-glycan antibodies in both serum and ascites. *Gynaekol. Geburtshilfliche Rundsch*

Presentations

10/2009	"Profiling of Anti-Glycan Antibodies reveals interesting findings in Epithelial Ovarian Cancer" Research Seminar, Carbohydrate Chemistry Lab, Russian Academy of Science, Moscow, Russian Federation
06/2009	"Serum Detection of Anti-Carbohydrate Antibodies in Serous Ovarian Cancer Patients" Annual Meeting of the Swiss Society for Gynecologists and Obstetricians (SGGG), Lugano, Switzerland
04/2009	"Profiling of Anti-Glycan Antibodies reveals interesting findings in Epithelial Ovarian Cancer" Biology PhD Retreat, Château de Bossey, Switzerland
09/2007	"Profiling of Anti-Glycan Autoantibodies as biomarker of malignancy status in Epithelial Ovarian Cancer" Cancer Network Zurich (CNZ) Cancer Biology PhD retreat, Eggberge, Switzerland
01/2007	"Sieben Jahre Functional Genomics beim Ovarialkarzinom: Was haben wir gelernt und welche Entwicklung zeichnet sich ab?" Oncology Seminar, Zurich, Switzerland

Awards

2009	Poster Price, Annual Meeting of the Swiss Society for Gynecologists and Obstetricians (SGGG), Lugano/Switzerland